

WEAK AND STRONG TYPE ESTIMATES FOR MAXIMAL TRUNCATIONS OF CALDERÓN-ZYGMUND OPERATORS ON A_p WEIGHTED SPACES

TUOMAS P. HYTÖNEN, MICHAEL T. LACEY, HENRI MARTIKAINEN, TUOMAS ORPONEN,
MARIA CARMEN REGUERA, ERIC T. SAWYER, AND IGNACIO URIARTE-TUERO

ABSTRACT. For $1 < p < \infty$, weight $w \in A_p$ and any L^2 -bounded Calderón-Zygmund operator T , we show that there is a constant $C_{T,p}$ so that we have the weak and strong type inequalities

$$\begin{aligned}\|T_{\natural}f\|_{L^{p,\infty}(w)} &\leq C_{T,p}\|w\|_{A_p}\|f\|_{L^p(w)} \\ \|T_{\natural}f\|_{L^p(w)} &\leq C_{T,p}\|w\|_{A_p}^{\max\{1,(p-1)^{-1}\}}\|f\|_{L^p(w)},\end{aligned}$$

where T_{\natural} denotes the maximal truncations of T , w is a weight, and $\|w\|_{A_p}$ denotes the Muckenhoupt A_p characteristic of w . These estimates are not improvable in the power of $\|w\|_{A_p}$. Our argument follows the outlines of the arguments of Lacey–Petermichl–Reguera (Math. Ann. 2010) and Hytönen–Pérez–Treil–Volberg (arXiv, 2010) with new ingredients, including a weak-type estimate for certain duals of T_{\natural} , and sufficient conditions for two weight inequalities in L^p for T_{\natural} . Our proof does not rely upon extrapolation.

CONTENTS

1. Overview and Introduction	2
2. Haar Shift Operators	5
3. Linearizing Maximal Operators	9
4. The Two Weight Estimates	11
5. Generalities of the Proof	13
6. Proof of the Weak-Type Inequality	14
7. First Steps in the Proof of the Strong Type Inequality	16
8. Strong-type estimates: the core	17
9. Unweighted weak-type (1,1) inequalities	21
10. Distributional Estimates	26
11. Verifying the Testing Conditions	33
12. Variations on the Main Theorem	37

Research of TPH and HM supported by the Academy of Finland grants 130166, 133264 and 218418.

Research of MTL, and MCR supported in part by NSF grant 0968499.

Research of TO supported by the Finnish Centre of Excellence in Analysis and Dynamics Research.

Research of ETS supported in part by NSERC.

Research of IU-T supported in part by the NSF, through grant DMS-0901524.

1. OVERVIEW AND INTRODUCTION

Our subject is weighted inequalities for maximal truncations T_{\sharp} of Calderón-Zygmund operators. There are two main results. First, we prove weak and strong norm estimates on T_{\sharp} on $L^p(w)$, that are sharp in the A_p characteristic of the weight w . In the generality of this paper, this was only known for the untruncated operators, a question investigated by many, culminating in the definitive result in [7].

Second, for dyadic Calderón-Zygmund operators, termed Haar Shift operators, we prove sufficient conditions for the weak and strong type two-weight inequalities T_{\sharp} . These estimates are effective in terms of a notion of complexity for the Haar shift, and while providing only sufficient conditions, are sharp enough in the A_p setting that we can conclude our first result from them.

We recall definitions.

Definition 1.1. A *Calderón-Zygmund operator* in \mathbb{R}^d is a *bounded* in L^2 integral operator with kernel $K(x, y)$, defined by the expression

$$\langle Tf, g \rangle = \iint f(x)g(y)K(x, y) dy$$

for all continuous compactly supported functions f, g with $\text{dist}(\text{supp}(f), \text{supp}(g)) > 0$. The kernel $K(x, y)$ satisfies the following growth and smoothness conditions for $x, x', y, y' \in \mathbb{R}^d$, $x \neq y$

$$|K(x, y)| \leq \frac{C_T}{|x - y|^d}, \quad x, y \in \mathbb{R}^d, x \neq y.$$

$$|K(x, y) - K(x', y)| + |K(y, x) - K(y, x')| \leq C_T \frac{|x - x'|^\alpha}{|x - y|^{d+\alpha}}, \quad |x - x'| < |x - y|/2.$$

Here, C_T is an absolute constant. We denote the maximal truncations of T by

$$T_{\sharp}f(x) := \sup_{0 < \epsilon < \nu} \left| \int_{\epsilon < |y| < \nu} f(y)K(x, y) dy \right|$$

It is well-known that T and T_{\sharp} extend to bounded operators on $L^p(\mathbb{R}^d)$, for $1 < p < \infty$.

Prominent examples include the Hilbert and Beurling transforms, as well as the vector R of Riesz transforms. If w is a weight on \mathbb{R}^d , namely a non-negative measure, with density also denoted as w that is non-negative almost everywhere, it is well-known [6] that R is bounded on $L^p(w)$, $1 < p < \infty$, if and only if w satisfies the famous Muckenhoupt A_p condition

$$\|w\|_{A_p} := \sup_Q |Q|^{-1} \int_Q w(dx) \left[|Q|^{-1} \int_Q \sigma(dx) \right]^{p-1}$$

where σ is the weight with density $w^{-1/(p-1)}$, which is dual to w . Note that $\|w\|_{A_p}$ is certainly not a norm.

On the other hand, determining the sharp dependence of Calderón-Zygmund operators on the quantity $\|w\|_{A_p}$ is not straight forward, as first pointed out by Buckley [2]. This direction has been intensively studied in recent years, with the sharp result for T established in [7], following the contributions of several. We refer the reader to the introductions of [4, 7, 11, 22] for more information about the history and range of techniques brought to bear on this problem.

Our first main result is this Theorem.

Theorem 1.2. *For T an $L^2(\mathbb{R}^d)$ bounded Calderón-Zygmund Operator,*

$$\begin{aligned} \|T_{\sharp}f\|_{L^{p,\infty}(w)} &\leq C_T\|w\|_{A_p}\|f\|_{L^p(w)}, \quad 1 < p < 2, \\ \|T_{\sharp}f\|_{L^p(w)} &\leq C_T\|w\|_{A_p}^{\max\{1,(p-1)^{-1}\}}\|f\|_{L^p(w)}, \quad 1 < p < \infty. \end{aligned}$$

Well known examples involving power weights (see the conclusion of [22]) show that all the estimates above are sharp. Indeed, these bounds match the best possible bounds for the untruncated operator T . The weak-type estimate was conjectured by Andrei Lerner [18], who also conjectured that the maximal truncations should have the same behavior in the A_p characteristic as the untruncated operators (personal communication). As far as we are aware, this is the first place in which the sharp estimates for T_{\sharp} have been established, and the weak-type inequality is new even for untruncated T .

We move to a discussion of the proof strategy for this Theorem. We will follow the outlines of the argument of [11], but the underlying details are substantially different. The strategy is summarized in Figure 1, and has the following points.

We begin with a Calderón-Zygmund Operator T , and the important step, identified in [7], is to write T as a rapidly convergent sum of *Haar Shift Operators* $\mathbb{S}_{m,n}$. See Definition 2.3, and Theorem 2.5.

Haar Shift Operators are themselves dyadic variants of Calderón-Zygmund Operators, and come with an essential notion of *complexity*, which is the measure of how many inter-related dyadic scales the operator has. As Calderón-Zygmund Operators, they satisfy many estimates already, but it is a vital point that in order to use the fact that T is a rapidly convergent sum of these operators, all relevant estimates must be shown to be *at most polynomial in complexity*. We will refer to this as an *effective estimate*. This requires that we revisit most facts about these operators, and verify that they meet this requirement.

The next crucial stage, the most complicated part of this argument, is to prove reasonably sharp *two weight inequalities* for Haar Shift Operators. The import here is that to prove our theorem, much of the argument must work in the generality of the two weight setting for a dyadic Calderón-Zygmund Operator. That the weight is in A_p is a fact that can only be used very sparingly. In this, we are following the pattern of [7–9, 12].

All of these prior works depended upon two-weight inequalities for the *untruncated* operator, and *only in L^2* . Here, we are concerned with two-weight inequalities for the *maximal truncations*; these estimates will apply in all L^p spaces, an important point as concerns the

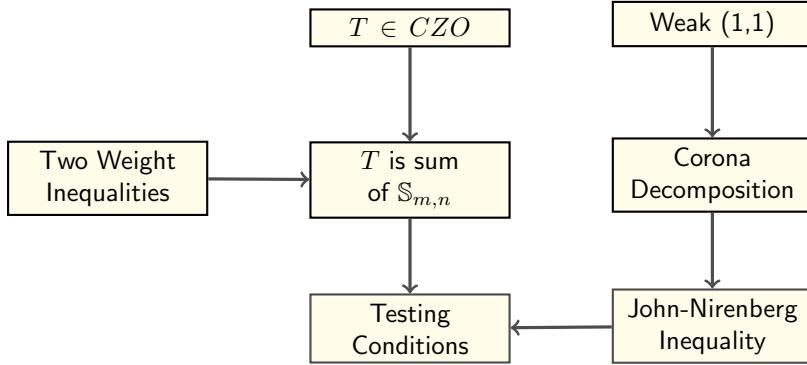


FIGURE 1. Proof Strategy

weak-type inequality. These estimates are taken up in §4, with the weak-type estimate being simple, and the strong type estimate being the most complicated estimate. Different variants of this argument have been used in [8, 9, 14, 15], with the point here being that the estimates in §4 track complexity. See this section for more history on these estimates.

The essential consequence of the two-weight inequalities is that they reduce the question of estimating the norm of T to that of testing the norm on a much simpler class of functions—weighted indicators of intervals. These conditions are in turn verified by using a chain of arguments that begins with the verification of certain weak- L^1 inequalities for the Haar Shift Operators. We need these weak-type bounds for the *adjoints of all linearizations* of the maximal truncation operator. This is an estimate not of a classical nature, and is taken up in §9.

This weak-integrability has a certain measure of uniformity. This permits the use of a *John-Nirenberg Inequality* that shows that uniform weak-integrability actually implies exponential integrability. This principle, again needed for certain maximal truncations, is formalized in §10.

In order to apply the John-Nirenberg Inequality, with the weight w fixed, we should decompose the collection of dyadic cubes into a *Corona Decomposition*. As we work with Haar Shifts, a decomposition of the cubes leads automatically to the decomposition of the operator \mathbb{S} . This leads to a decomposition of $\mathbb{S}(w\mathbf{1}_E)$ into terms which are individually very nicely behaved.

Finally, the testing conditions can be verified, and using the exponential integrability from the Corona Decomposition, one can give a simple verification of these conditions. This part of the argument is new to this paper. This argument will *not* appeal to extrapolation, a common technique in this subject. Indeed, the weak type estimate we prove does not seem to lend itself to extrapolation.

In the ultimate section, we provide some variations and consequences of Theorem 1.2.

2. HAAR SHIFT OPERATORS

In this section, we introduce fundamental dyadic approximations of Calderón–Zygmund operators, the Haar shifts, and make a detailed reduction of the Main Theorem 1.2 to a similar statement, Theorem 2.10, in this dyadic model.

Definition 2.1. A *dyadic grid* is a collection \mathcal{D} of cubes so that for each Q we have that

- (1) The set of cubes $\{Q' \in \mathcal{D} : |Q'| = |Q|\}$ partition \mathbb{R}^d , ignoring overlapping boundaries of cubes.
- (2) Q is a union of cubes in a collection $\text{Child}(Q) \subset \mathcal{D}$, called the *children of Q* . There are 2^d children of Q , each of volume $|Q'| = 2^{-d}|Q|$.

We refer to any subset of a dyadic grid as simply a *grid*.

The standard choice for \mathcal{D} consists of the cubes $2^k \prod_{s=1}^d [n_s, n_s + 1)$ for $k, n_1, \dots, n_d \in \mathbb{Z}$. But, the main result of this section, Theorem 2.5, depends upon a random family of dyadic grids.

In higher dimensions, we mention that the martingale differences are finite rank projections, but there is no canonical choice of the Haar functions in this case. We make the following definition.

Definition 2.2. Let Q be a dyadic cube, a *generalized Haar function* associated to Q , h_Q , is a linear combination of the indicator functions $\{\mathbf{1}_Q\} \cup \{\mathbf{1}_{Q'} : Q' \text{ is a child of } Q\}$,

$$h_Q = \sum_{Q' \in \text{Child}(Q)} c_{Q'} \mathbf{1}_{Q'}$$

We say h_Q is a *Haar function* if in addition $\int h_Q = 0$, that is, a Haar function is orthogonal to constants on its support.

Definition 2.3. For integers $(m, n) \in \mathbb{Z}_+^2$, we say that a linear operator \mathbb{S} is a *(generalized) Haar shift operator of complexity type (m, n)* if

$$\mathbb{S}f(x) = \sum_{Q \in \mathcal{D}} \mathbb{S}_Q f(x) = \sum_{Q \in \mathcal{D}} \sum_{\substack{Q', R' \in \mathcal{D} \\ Q', R' \subset Q}}^{(m, n)} \frac{\langle f, h_{R'}^{Q'} \rangle}{|Q|} k_{Q'}^{R'}(x)$$

where here and throughout $\ell(Q) = |Q|^{1/d}$, and

- in the second sum, the superscript $^{(m, n)}$ on the sum means that in addition we require $\ell(Q') = 2^{-m} \ell(Q)$ and $\ell(R') = 2^{-n} \ell(Q)$, and
- the function $h_{R'}^{Q'}$ is a (generalized) Haar function on R' , and $k_{Q'}^{R'}$ is one on Q' , with the normalization that

$$\|h_{R'}^{Q'}\|_\infty \leq 1, \quad \|k_{Q'}^{R'}\|_\infty \leq 1.$$

Here, and throughout the paper, $\ell(Q) = |Q|^{1/d}$ is the side length of the cube Q . We say that the *complexity* of \mathbb{S} is $\kappa := \max(m, n, 1)$.

A generalized Haar shift thus has the form

$$\mathbb{S}f(x) = \sum_{Q \in \mathcal{D}} \frac{1}{|Q|} \int_Q s_Q(x, y) f(y) dy = \int_{\mathbb{R}^n} K_{\mathbb{S}}(x, y) f(y) dy,$$

where s_Q , the kernel of the component \mathbb{S}_Q , is supported on $Q \times Q$ and $\|s_Q\|_{\infty} \leq 1$. It is easy to check that

$$|K_{\mathbb{S}}(x, y)| \lesssim \frac{1}{|x - y|^d}.$$

The Haar shifts are automatically bounded on L^2 with $\|\mathbb{S}f\|_{L^2} \leq \|f\|_{L^2}$. This follows from the imposed normalizations and simple orthogonality considerations. It is clear that all *restricted shifts* $\mathbb{S}_{\mathcal{Q}}f = \sum_{Q \in \mathcal{Q}} \mathbb{S}_Q f$ are also Haar shifts, and hence uniformly bounded on L^2 for any $\mathcal{Q} \subset \mathcal{D}$. We will extensively exploit these restricted shifts in the argument.

The generalized Haar shifts are only of relevance to us in two particular special cases of complexity type $(0, 0)$, where $\mathbb{S}_Q f = |Q|^{-1} \langle f, h_Q^Q \rangle k_Q^Q = |Q|^{-1} \langle f, h_Q \rangle k_Q$. These are the *paraproduct*, where $h_Q = \mathbf{1}_Q$ and k_Q is a Haar function for all Q , and the *dual paraproduct*, where h_Q is a Haar function and $k_Q = \mathbf{1}_Q$ for all Q .

It is well-known that the (normalized) L^2 boundedness of a (dual) paraproduct is equivalent to the Carleson condition

$$\sum_{Q \subset R} \|k_Q\|_{L^2}^2 \leq |R| \quad \left(\sum_{Q \subset R} \|h_Q\|_{L^2}^2 \leq |R| \right) \quad \forall R \in \mathcal{D}.$$

These conditions are also uniformly inherited by all restricted (dual) paraproducts $\mathbb{S}_{\mathcal{Q}}$.

Note that for both Haar shifts and the paraproduct, we have

$$\int \mathbb{S}_Q f = 0 \quad \forall Q \in \mathcal{D},$$

an important cancellation property in some of the later arguments. This is not the case for the dual paraproduct, for which a separate case study is needed at some points.

Remark 2.4. Let

$$\delta(x, y) = \min \{ \ell(Q) : x, y \in Q \in \mathcal{D} \}$$

be the *dyadic distance* between x and y . The kernel $K_{\mathbb{S}}(x, y)$ of \mathbb{S} satisfies the size and smoothness conditions for a dyadic Calderón–Zygmund kernel:

$$\begin{aligned} |K_{\mathbb{S}}(x, y)| &\leq \frac{2}{\delta(x, y)^d}, \\ |K_{\mathbb{S}}(x, y) - K_{\mathbb{S}}(x', y)| &= 0 \text{ if } \frac{\delta(x, x')}{\delta(x, y)} < \frac{1}{2^m}, \\ |K_{\mathbb{S}}(x, y) - K_{\mathbb{S}}(x, y')| &= 0 \text{ if } \frac{\delta(y, y')}{\delta(x, y)} < \frac{1}{2^n}. \end{aligned}$$

This is a more general dyadic kernel condition than one studied in [1], called *perfect dyadic*, which corresponds to $m = n = 0$ in our framework.

The relevance of Haar shifts to Classical Analysis is explained by the following Theorem, one of the main results of [11] (see [11, Theorem 4.1]; also [7, Theorem 4.2]). This Theorem must be formulated in terms of a *random* dyadic grids. But the nature of this construction of grids is immaterial to the arguments of this paper, and refer the reader to these references for proofs, history, and further discussion of this result.

Theorem 2.5. *There is a collection of random dyadic grids $\{\mathcal{D}_\beta : \beta \in \mathcal{B}\}$, with expectation operator \mathbb{E}_β , for which the following holds. Let T be a Calderón-Zygmund Operator T with smoothness parameter δ . Then, for all bounded and compactly supported functions f and g , we can write*

$$\langle Tf, g \rangle = C \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} \langle \mathbb{S}_{m,n}^\beta f, g \rangle$$

where

- $\mathbb{S}_{m,n}^\beta$ is a Haar shift of complexity type (m, n) for all $(m, n) \in \mathbb{Z}_+^2 \setminus \{(0, 0)\}$;
- $\mathbb{S}_{0,0}^\beta$ is the sum of a Haar shift of type $(0, 0)$, a paraproduct, and a dual paraproduct;
- the constant C is a function of T , and of the smoothness parameter δ .

In particular we have the uniform estimate $\|\mathbb{S}_{m,n}^\beta\|_{L^2 \rightarrow L^2} \leq 1$.

We define the maximal truncations of a Haar Shift as follows.

Definition 2.6. Suppose that \mathbb{S} is a generalized Haar shift. Define the associated *maximal truncations* by

$$\mathbb{S}_\sharp f(x) \equiv \sup_{0 < \epsilon \leq v < \infty} |\mathbb{S}_{\epsilon,v} f(x)|,$$

$$(2.7) \quad \mathbb{S}_{\epsilon,v} f \equiv \sum_{Q \in \mathcal{D} : \epsilon \leq \ell(Q) \leq v} \mathbb{S}_Q f(x).$$

Proposition 2.8. *We have the pointwise bound*

$$T_\sharp f(x) \leq C(T) \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} (\mathbb{S}_{m,n}^\beta)_\sharp f(x) + C(T) M f(x),$$

where M is the Hardy-Littlewood maximal operator.

Proof. Theorem 2.5 says that

$$\langle Tf, g \rangle = C(T) \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} \langle \mathbb{S}_{m,n}^\beta f, g \rangle$$

for bounded and compactly supported functions f and g . By choosing

$$f = \frac{\mathbf{1}_{B(y,\epsilon)}}{|B(y, \epsilon)|} \quad \text{and} \quad g = \frac{\mathbf{1}_{B(x,\epsilon)}}{|B(x, \epsilon)|}$$

and taking the limit as $\epsilon \rightarrow 0$, dominated convergence implies the pointwise identity

$$K(x, y) = C(T) \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} \sum_{Q \in \mathcal{D}^\beta} \frac{s_Q^{m,n}(x, y)}{|Q|}.$$

For $\epsilon > 0$, this implies by Fubini's theorem that

$$\begin{aligned} & \int_{|x-y| \geq \epsilon} K(x, y) f(y) dy \\ &= C(T) \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} \sum_{Q \in \mathcal{D}^\beta} \frac{1}{|Q|} \int_{B(x, \epsilon)^c} s_Q^{m,n}(x, y) f(y) dy. \end{aligned}$$

Let us then decompose

$$\begin{aligned} \sum_{Q \in \mathcal{D}^\beta} \frac{1}{|Q|} \int_{B(x, \epsilon)^c} s_Q^{m,n}(x, y) f(y) dy &= \sum_{\substack{Q \in \mathcal{D}^\beta \\ \ell(Q) > \epsilon'}} \frac{1}{|Q|} \int s_Q^{m,n}(x, y) f(y) dy \\ &\quad - \sum_{\substack{Q \in \mathcal{D}^\beta \\ \ell(Q) > \epsilon'}} \frac{1}{|Q|} \int_{B(x, \epsilon)} s_Q^{m,n}(x, y) f(y) dy \\ &\quad + \sum_{\substack{Q \in \mathcal{D}^\beta \\ \ell(Q) \leq \epsilon'}} \frac{1}{|Q|} \int_{B(x, \epsilon)^c} s_Q^{m,n}(x, y) f(y) dy \\ &= I + II + III, \end{aligned}$$

where $\epsilon' = \epsilon/(2\sqrt{d})$. By definition, there holds $|I| \leq (\mathbb{S}_{m,n}^\beta)_{\sharp} f(x)$. There also holds

$$|II| \leq \sum_{k: 2^k > \epsilon'} 2^{-kd} \int_{B(x, \epsilon)} |f| \lesssim \epsilon^d \left(\sum_{k: 2^k > \epsilon'} 2^{-kd} \right) Mf(x) \lesssim Mf(x).$$

Furthermore, we actually have $III = 0$, since there we must have $|x - y| \leq d(Q) = \sqrt{d}\ell(Q) \leq \sqrt{d}\epsilon' = \epsilon/2 < \epsilon$. We have thus shown that

$$T_\epsilon f(x) \leq C(T) \mathbb{E}_\beta \sum_{(m,n) \in \mathbb{Z}_+^2} 2^{-(m+n)\delta/2} (\mathbb{S}_{m,n}^\beta)_{\sharp} f(x) + C(T) Mf(x)$$

for every $\epsilon > 0$, and from this the proposition follows. \square

Proposition 2.8 and Buckley's [2] sharp weighted bounds for the maximal operator,

$$(2.9) \quad \|Mf\|_{L^{p,\infty}(w)} \lesssim \|w\|_{A_p}^{1/p} \|f\|_{L^p(w)}, \quad \|Mf\|_{L^p(w)} \lesssim \|w\|_{A_p}^{1/(p-1)} \|f\|_{L^p(w)},$$

reduce the proof of the Main Theorem 1.2 to the verification of the following dyadic variant, a task which occupies the rest of this paper.

Theorem 2.10. *Let \mathbb{S} be a Haar shift operator with complexity κ , a paraproduct, or a dual paraproduct. For $1 < p < \infty$ and $w \in A_p$, we then have the estimates*

$$(2.11) \quad \|\mathbb{S}_\sharp f\|_{L^{p,\infty}(w)} \lesssim \kappa \|w\|_{A_p} \|f\|_{L^p(w)},$$

$$(2.12) \quad \|\mathbb{S}_\sharp f\|_{L^p(w)} \lesssim \kappa \|w\|_{A_p}^{\max\{1,(p-1)^{-1}\}} \|f\|_{L^p(w)}$$

Indeed, any polynomial dependence on the complexity parameter κ would suffice for Theorem 1.2, but a careful tracing of the constants will even provide the *linear* dependence, as stated. Even for the untruncated shifts \mathbb{S} in $L^2(w)$, this improves on the quadratic in κ bound established in [11] (but we are not aware of an application where this precision in the dependence on κ would be of importance).

The dependence on κ and the weight constant $\|w\|_{A_p}$ arises from the following points of the proof below: First, we establish *two-weight inequalities* of the form

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^{p,\infty}(w)} \lesssim \{\kappa \mathfrak{M}_{p,\text{weak}} + \mathfrak{T}_p\} \|f\|_{L^p(\sigma)},$$

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)} \lesssim \{\kappa \mathfrak{M}_p + \mathfrak{T}_p + \mathfrak{N}_p\} \|f\|_{L^p(\sigma)},$$

where $\mathfrak{M}_{p,\text{weak}}$ and \mathfrak{M}_p are the best constants from certain maximal inequalities, while \mathfrak{T}_p and \mathfrak{N}_p are the best constants from appropriate *testing conditions* for the operator \mathbb{S}_\sharp . Here, w and σ are allowed to be an arbitrary pair of weights, with no relation to each other.

Second, we specialize to the one-weight situation with $\sigma = w^{1-p'}$, using a well-known dual-weight formulation of the bounds to be proven, (2.11) and (2.12). We need to estimate the above four constants in this situation. The maximal constants are independent of \mathbb{S}_\sharp and thus of κ , and they satisfy $\mathfrak{M}_{p,\text{weak}} \lesssim \|w\|_{A_p}^{1/p}$ and $\mathfrak{M}_p \lesssim \|w\|_{A_p}^{1/(p-1)}$ by the sharp maximal function inequalities of Buckley (2.9). For the two testing constants related to \mathbb{S}_\sharp , we obtain the linear in κ bounds

$$\mathfrak{T}_p \lesssim \kappa \|w\|_{A_p}, \quad \mathfrak{N}_p \lesssim \kappa \|w\|_{A_p}^{1/(p-1)}.$$

This dependence comes from the fact that the proof of the John–Nirenberg style estimates of (10.13) requires *separating the scales* of \mathbb{S} by dividing it into $\kappa + 1$ parts, each of which contains nonzero components \mathbb{S}_Q only for a fixed value of $\log_2 \ell(Q) \bmod (\kappa + 1)$. For these separated parts of \mathbb{S} , our bounds will be independent of κ , and it remains to sum up.

3. LINEARIZING MAXIMAL OPERATORS

A fundamental tool is derived from (the usual) general maximal function estimates that hold for any measure. In particular, for weight w we define

$$M_w f(x) \equiv \sup_{Q \in \mathcal{D}} \mathbf{1}_Q(x) \mathbb{E}_Q^w |f|,$$

$$\mathbb{E}_Q^w f \equiv w(Q)^{-1} \int_Q f w(dx).$$

The notation $\mathbb{E}_Q f$ means that the implied measure is Lebesgue. It is a basic fact, proved by exactly the same methods that prove the non-weighted inequality, that we have the estimate below, which will be used repeatedly.

Theorem 3.1. *We have the inequalities*

$$(3.2) \quad \|M_w f\|_{L^p(w)} \lesssim \|f\|_{L^p(w)}, \quad 1 < p < \infty.$$

We use the method of linearizing maximal operators. This is familiar in the context of the maximal function, and we make a comment about it here. Let $\{E(Q) : Q \in \mathcal{D}\}$ be any selection of measurable disjoint sets $E(Q) \subset Q$ indexed by the dyadic cubes. Define a corresponding linear operator N by

$$N\phi \equiv \sum_{Q \in \mathcal{D}} \mathbf{1}_{E(Q)} \mathbb{E}_Q^w f.$$

Then, the universal Maximal function bound (3.2) is equivalent to the bound $\|Nf\|_{L^p(w)} \lesssim \|f\|_{L^p(w)}$ with implied constant independent of w and the sets $\{E(Q) : Q \in \mathcal{D}\}$. This estimate will be used repeatedly below.

There is a related way to linearize \mathbb{S}_\natural , which deserves careful comment as we would like, at different points, to treat \mathbb{S}_\natural as a linear operator. While it is not a linear operator, \mathbb{S}_\natural is a pointwise supremum of the linear truncation operators $\mathbb{S}_{\epsilon,v}$, and as such, the supremum can be linearized with measurable selection of the truncation parameters.

Definition 3.3. We say that \mathbb{L} is a linearization of \mathbb{S}_\natural if there are measurable functions $\epsilon(x), v(x) \in (0, \infty)$ and $\vartheta(x) \in [0, 2\pi)$ such that, using (2.7), we have

$$\mathbb{L}f(x) = e^{i\vartheta(x)} \mathbb{S}_{\epsilon(x), v(x)} f(x) \geq 0, \quad x \in \mathbb{R}^d.$$

Note that the requirement $\mathbb{L}f(x) \geq 0$ defines $\vartheta(x)$ everywhere except when $\mathbb{S}_{\epsilon(x), v(x)} f(x) = 0$. Also, for fixed f we can always choose a linearization L so that $\mathbb{S}_\natural f(x) \leq 2\mathbb{L}f(x)$ for all x .

A key advantage of \mathbb{L} is that it is a linear operator, and as such it has an adjoint, given by the formal expressions

$$(3.4) \quad \begin{aligned} \mathbb{L}^* \nu(y) &= \sum_{Q \in \mathcal{D}} \mathbb{S}_Q^* (\mathbf{1}_{\{\epsilon(\cdot) \leq \ell(Q) \leq v(\cdot)\}} e^{i\vartheta(\cdot)} \nu)(y) \\ &= \sum_{Q \in \mathcal{D}} \frac{1}{|Q|} \int_Q s_Q(x, y) \mathbf{1}_{\{\epsilon(x) \leq \ell(Q) \leq v(x)\}} e^{i\vartheta(x)} \nu(dx). \end{aligned}$$

The following ‘smoothness’ property of \mathbb{L}^* is an important observation in the proof of our two weight estimates.

Lemma 3.5. *Suppose that for a measure ν and cube Q_0 we have $|\nu|(Q_0) = 0$. Suppose that \mathbb{S} has complexity type (m, n) . Then $\mathbb{L}^* \nu(\cdot)$ is constant on subcubes $Q' \subset Q_0$ with $\ell(Q') \leq 2^{-n} \ell(Q_0)$.*

Proof. For $y \in Q_0$, the sum in (3.4) defining the adjoint operator becomes

$$\mathbb{L}^* \nu(y) = \sum_{Q \in \mathcal{D}: Q \supseteq Q_0} \frac{1}{|Q|} \int_Q s_Q(x, y) \mathbf{1}_{\{\epsilon(x) \leq \ell(Q) \leq v(x)\}} e^{i\vartheta(x)} \nu(dx).$$

As a function of y , the kernel $s_Q(x, y)$ is constant on the subcubes of $R' \subset Q$ with $\ell(R') \leq 2^{-n-1}\ell(Q)$. Thus for $Q \supsetneq Q_0$, it is in particular constant on the subcubes $Q' \subset Q_0$ with $\ell(Q') \leq 2^{-n}\ell(Q_0)$. \square

4. THE TWO WEIGHT ESTIMATES

We are interested in tracking complexity dependence in two weight inequalities for Haar Shift Operators, as defined in §2. We study the maximal truncations of such operators, and obtain sufficient conditions for the weak and strong type (p, p) two weight inequalities for such operators. Our main results are Theorem 4.3 for the weak type result, and Theorem 4.7 for the strong type result. These Theorems give sufficient conditions in terms of the Maximal Function, and certain *testing conditions*. Of particular import here is that these sufficient conditions are efficient in terms of the *complexity* of the Haar Shift operator.

Our primary focus concerns extensions of the dyadic $T1$ Theorem to the two weight setting. These considerations are motivated in part by a well developed theory of two weight estimates for positive operators. These Theorems have formulations strikingly similar to the $T1$ Theorem, which theory encompasses the Theorems due to one of us concerning two weight, both strong and weak type, for the maximal operator [23] and fractional integral operators [24], [25]. There is also the bilinear embedding inequality of Nazarov-Treil-Volberg [21]. We refer the reader to [13] for a discussion of these results.

There is a beautiful result of Nazarov-Treil-Volberg [20], a two-weight version of the $T1$ theorem. A subcase of their result was proved for Haar Shifts, with an effective bound on complexity in [11, Theorem 3.4].

Theorem 4.1. *Let \mathbb{S} be a Haar Shift operator of complexity κ , as in Definition 2.3. Let σ, w be two positive locally finite measures. We have the inequality*

$$\|\mathbb{S}(f\sigma)\|_{L^2(w)} \lesssim \kappa\{\mathfrak{S} + \mathfrak{S}^*\} + \kappa^2\|w, \sigma\|_{A_2}^{1/2}$$

where the three quantities above are defined by

$$\begin{aligned} \|w, \sigma\|_{A_2} &:= \sup_Q \frac{w(Q)}{|Q|} \frac{\sigma(Q)}{|Q|} \\ \mathfrak{S} &:= \sup_Q \sigma(Q)^{-1/2} \|\mathbf{1}_Q \mathbb{S}(\sigma \mathbf{1}_Q)\|_{L^2(w)} . \\ (4.2) \quad \mathfrak{S}^* &:= \sup_Q w(Q)^{-1/2} \|\mathbf{1}_Q \mathbb{S}^*(w \mathbf{1}_Q)\|_{L^2(\sigma)} . \end{aligned}$$

The first of the three conditions is the two weight A_2 condition; the remaining two are the testing conditions. The proof is fundamentally restricted to the case of $p = 2$, nor does it address maximal truncations. We will consider the case of $1 < p < \infty$ and obtain sufficient conditions for the two weight inequalities for the maximal operator \mathbb{S}_\sharp . First we give the weak type result. Below, M denotes the maximal function.

Theorem 4.3. *Let \mathbb{S} be a generalized Haar shift of complexity κ as in Definition 2.3. Then we have the weak type inequality*

$$(4.4) \quad \|\mathbb{S}_\sharp(f\sigma)\|_{L^{p,\infty}(w)} \lesssim (\kappa\mathfrak{M}_{p,\text{weak}} + \mathfrak{T}_p) \|f\|_{L^p(\sigma)},$$

where the constants $\mathfrak{M}_{p,\text{weak}}$ and \mathfrak{T}_p are the best such in the following inequalities

$$(4.5) \quad \begin{aligned} \|M(f\sigma)\|_{L^{p,\infty}(w)} &\leq \mathfrak{M}_{p,\text{weak}} \|f\|_{L^p(\sigma)}, \\ \int_Q \mathbb{S}_\sharp(\sigma f \mathbf{1}_Q) w(dx) &\leq \mathfrak{T}_p \|f\|_{L^p(\sigma)} w(Q)^{1/p'}. \end{aligned}$$

The point of this Theorem is that to check the weak-type inequality for \mathbb{S}_\sharp , it suffices to check the weak-type inequality for the simpler maximal operator M , and to check only particular instances of the weak-type inequality for \mathbb{S}_\sharp . It is also important that the complexity κ appears with polynomial growth.

The dual testing condition (4.5) looks rather complicated, with the appearance of $f \in L^p(\sigma)$ in it. However, \mathbb{S}_\sharp appears to just the first power, and it is a close relative of (4.2). Indeed, (4.5) has a more convincing formulation in the linearizations. It is equivalent to the dual testing condition

$$(4.6) \quad \|\mathbf{1}_Q \mathbb{L}^*(\mathbf{1}_Q g w)\|_{L^{p'}(\sigma)} \leq \mathfrak{T}_p w(Q)^{1/p'} \|g\|_\infty,$$

This holds uniformly over all choices of linearizations, which fact is referred to repeatedly below. Inequality (4.6) reflects the fact that the dual of a weak type inequality is a restricted strong type inequality.

Our strong type result will require duals \mathbb{L}^* of linearizations \mathbb{L} of \mathbb{S}_\sharp in order to state the nonstandard testing condition in (4.8). These were defined in Section 3 above.

Theorem 4.7. *Let \mathbb{S} be a generalized Haar shift of complexity κ as in Definition 2.3. We have the following quantitative estimate:*

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)} \lesssim \{\kappa\mathfrak{M}_p + \mathfrak{T}_p + \mathfrak{N}_p\} \|f\|_{L^p(\sigma)},$$

where \mathfrak{T}_p is defined in (4.5), and the numbers \mathfrak{M}_p and \mathfrak{N}_p are defined in

$$\|M(f\sigma)\|_{L^p(w)} \leq \mathfrak{M}_p \|f\|_{L^p(\sigma)},$$

$$(4.8) \quad \mathfrak{N}_p^p \equiv \sup_{\|\varphi\|_\infty \leq 1} \sup_{Q_0} \frac{1}{\sigma(Q_0)} \int_{Q_0} \sup_{Q \subset Q_0} \mathbf{1}_Q \left(\frac{1}{w(Q)} \int_Q |\mathbb{L}^*(\mathbf{1}_Q \varphi w)(y)| \sigma(dy) \right)^p w(dx).$$

As a new kind of complication compared to the weak-type case, we have the nonstandard testing condition (4.8). Its primary difficulty is the appearance of $|\mathbb{L}^*|$ integrated over Q with respect to σ , but then divided by $w(Q)$ rather than the usual $\sigma(Q)$. Also there is an additional supremum, with the argument of \mathbb{L}^* dependent upon the cube Q over which we are taking the supremum.

The method of proof is an extension of that of Sawyer's approach to the two weight fractional integrals [25], but see also [13]. This argument follows the outlines of the proof in [14], which proves variants of Theorem 4.3 and Theorem 4.7 for smooth Calderón-Zygmund operators. The current arguments are, naturally, much easier while retaining the essential ideas and techniques of [14]. (The reader can also compare the arguments of this paper to those of [13].)

Let us give a guide to the next few sections of this paper, which are concerned with the proof of the above two-weight results.

- §5: Collects facts central to the proofs, maximal functions, linearizations of maximal functions, Whitney decompositions, and an important maximum principle.
- §6: The weak-type result Theorem 4.3 is proved.
- §7: Sufficient conditions for the strong type result are stated; the classical part of the proof of the strong type result Theorem 4.7 is begun.
- §8: This section contains the core of the proof of Theorem 4.7.

5. GENERALITIES OF THE PROOF

5.1. Whitney Decompositions. We make general remarks about the sets

$$(5.1) \quad \Omega_k = \{\mathbb{S}_\sharp(f\sigma) > 2^k\}$$

where f is a finite linear combination of indicators of dyadic cubes. For points x sufficiently far away from the support of f , we will have that $\mathbb{S}_\sharp(f\sigma)$ is dominated by the maximal function $M(f\sigma)$. Hence, the sets Ω_k will be open with compact closure.

Let $Q^{(1)}$ denote the dyadic parent of Q , and inductively define $Q^{(j+1)} = (Q^{(j)})^{(1)}$. For a nonnegative integer ζ , let \mathcal{Q}_k be the collection of maximal dyadic cubes Q such that $Q^{(\zeta)} \subset \Omega_k$. Then

$$\Omega_k = \bigcup_{Q \in \mathcal{Q}_k} Q \quad (\text{disjoint cover}),$$

$$Q^{(\zeta)} \subset \Omega_k, \quad Q^{(\zeta+1)} \cap \Omega_k^c \neq \emptyset \quad (\text{Whitney condition,})$$

$$Q \in \mathcal{Q}_k, \quad Q' \in \mathcal{Q}_l, \quad Q \subsetneq Q' \quad \text{implies} \quad k > l \quad (\text{nested property}).$$

Remark 5.2. In the proof of the weak type theorem we will take $\zeta = 0$. In the proof of the strong type theorem we will take $\zeta = n + 1 \leq \kappa + 1$, when the shift under consideration has complexity type (m, n) .

5.2. Maximum Principle. A fundamental tool is the use of what we term here as a 'maximum principle' (we could also use the term 'good- λ technique'): Subject to the assumption that the maximal function M is of small size, we will be able to see that the maximal truncations are large due to the restriction of the function to a local cube. This leads to an essential 'localization' of the singular integrals.

Theorem 5.3 (Maximum Principle). *Let \mathbb{S} be a generalized Haar shift of complexity type (m, n) . For any cube $Q \in \mathcal{Q}_k$ as in the Whitney decomposition of Ω_k in (5.1) above with parameter ζ , we have the pointwise inequality*

$$(5.4) \quad \begin{aligned} \mathbb{S}_\sharp(f\sigma)(x) &\leq \sup_{\epsilon \leq v \leq \ell(Q)} |\mathbb{S}_{\epsilon, v}(f\sigma)(x)| + 2^k + (\zeta + n + 1)M(f\sigma)(x) \\ &\leq \mathbb{S}_\sharp(f1_Q\sigma)(x) + 2^k + (\zeta + n + 1)M(f\sigma)(x). \end{aligned}$$

There is a corresponding Maximum Principle in Section 3.3 of [13], which is very effective in the positive operator case. As our operators are not positive, and as we are ultimately only interested in the one weight situation, we have the maximal function M on the right in (5.4).

Proof. Note that the second inequality of the claim is obvious, since $\mathbb{S}_{\epsilon, v}(f\sigma)(x) = \mathbb{S}_{\epsilon, v}(f1_Q\sigma)(x)$ when $x \in Q$ and $\epsilon \leq v \leq \ell(Q)$. We prove the first inequality.

Let $x \in Q \in \mathcal{Q}_k$, and let $r := (\zeta + 1) + (n + 1)$. Then

$$\begin{aligned} \mathbb{S}_{\epsilon, v}(f\sigma)(x) &= \sum_{I: \epsilon \leq \ell(I) \leq v} \mathbb{S}_I(f\sigma)(x) \\ &= \left(\sum_{\substack{I: \epsilon \leq \ell(I) \leq v \\ \ell(I) \leq \ell(Q)}} + \sum_{\substack{I: \epsilon \leq \ell(I) \leq v \\ \ell(I) \geq 2^r \ell(Q)}} + \sum_{\substack{I: \epsilon \leq \ell(I) \leq v \\ \ell(Q) < \ell(I) < 2^r \ell(Q)}} \right) \mathbb{S}_I(f\sigma)(x) \\ &= \mathbb{S}_{\epsilon, \min(v, \ell(Q))}(f\sigma)(x) + \mathbb{S}_{\max(\epsilon, 2^r \ell(Q)), v}(f\sigma)(x) + \sum_{\substack{I: \epsilon \leq \ell(I) \leq v \\ \ell(Q) < \ell(I) < 2^r \ell(Q)}} \mathbb{S}_I(f\sigma)(x). \end{aligned}$$

The first term on the right is clearly dominated by $\sup_{\epsilon' \leq v' \leq \ell(Q)} |\mathbb{S}_{\epsilon', v'}(f\sigma)(x)|$. All $\mathbb{S}_I(f\sigma)$ participating in the second term have $\ell(I) \geq 2^{n+1} \ell(Q^{(\zeta+1)})$, so they are constant on $Q^{(\zeta+1)}$. Hence we may replace x by some $\bar{x} \in Q^{(\zeta+1)} \setminus \Omega_k$ (which is nonempty by definition of \mathcal{Q}_k). So this second term is dominated by

$$\mathbb{S}_\sharp(f\sigma)(\bar{x}) \leq 2^k.$$

Finally, the last term contains at most $r - 1 = \zeta + n + 1$ summands, each of which is dominated by $M(f\sigma)(x)$. \square

6. PROOF OF THE WEAK-TYPE INEQUALITY

We prove Theorem 4.3, stating that

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^{p, \infty}(w)} \lesssim (\kappa \mathfrak{M}_{p, \text{weak}} + \mathfrak{T}_p) \|f\|_{L^p(\sigma)}.$$

To this end, we need to estimate the quantities

$$w(\mathbb{S}_\sharp(f\sigma) > 4 \cdot 2^k) \leq w(M(f\sigma) > \eta 2^k) + w(\mathbb{S}_\sharp(f\sigma) > 4 \cdot 2^k, M(f\sigma) \leq \eta 2^k),$$

where the small parameter η is to be chosen shortly. Since $\Omega_{k+2} \subset \Omega_k = \bigcup_{Q \in \mathcal{Q}_k} Q$, where we take the Whitney decomposition with parameter $\zeta = 0$, we further have that

$$\begin{aligned} w(\mathbb{S}_\sharp(f\sigma) > 4 \cdot 2^k, M(f\sigma) \leq \eta 2^k) \\ &= \sum_{Q \in \mathcal{Q}_k} w(Q \cap \{\mathbb{S}_\sharp(f\sigma) > 4 \cdot 2^k, M(f\sigma) \leq \eta 2^k\}) =: \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)). \end{aligned}$$

On $E_k(Q) \subset Q$, the maximum principle gives that

$$\begin{aligned} 4 \cdot 2^k < \mathbb{S}_\sharp(f\sigma) &\leq \mathbb{S}_\sharp(f1_Q\sigma) + 2^k + (2n+3)M(f\sigma) \\ &\leq \mathbb{S}_\sharp(f1_Q\sigma) + 2^k + (2n+3)\eta 2^k \leq \mathbb{S}_\sharp(f1_Q\sigma) + 2 \cdot 2^k, \end{aligned}$$

provided that $\eta := (2\kappa+3)^{-1} \leq (2n+3)^{-1}$. Thus

$$\mathbb{S}_\sharp(f\sigma 1_Q) > 2^{k+1} \quad \text{on } E_k(Q).$$

Putting these considerations together, we obtain (for another small parameter $\delta > 0$)

$$\begin{aligned} (4 \cdot 2^k)^p w(\mathbb{S}_\sharp(f\sigma) > 4 \cdot 2^k) \\ &\leq 4^p 2^{kp} w(M(f\sigma) > \eta 2^k) + 2^p \sum_{Q \in \mathcal{Q}_k} 2^{(k+1)p} w(E_k(Q)) \\ &\leq 4^p \eta^{-p} \|M(f\sigma)\|_{L^{p,\infty}(w)}^p + 4^p 2^{kp} \sum_{\substack{Q \in \mathcal{Q}_k \\ w(E_k(Q)) < \delta w(Q)}} \delta w(Q) \\ &\quad + 2^p \sum_{\substack{Q \in \mathcal{Q}_k \\ w(E_k(Q)) \geq \delta w(Q)}} w(E_k(Q)) \left(\frac{1}{w(E_k(Q))} \int_{E_k(Q)} \mathbb{S}_\sharp(f1_Q\sigma) w \right)^p \\ &\leq 2^p \eta^{-p} \mathfrak{M}_{p,\text{weak}}^p \|f\|_{L^p(\sigma)}^p + 4^p \delta 2^{kp} w(\mathbb{S}_\sharp(f\sigma) > 2^k) \\ &\quad + 2^p \sum_{Q \in \mathcal{Q}_k} \delta^{1-p} w(Q)^{1-p} \left(\int_Q \mathbb{S}_\sharp(f1_Q\sigma) w \right)^p \end{aligned}$$

Picking a k for which the left side is close to its supremum, and choosing $\delta = 10^{-p}$, we can absorb the middle term on the right to the left side. Recalling the choice of η and using the testing condition to the integrals in the last term, we obtain

$$\begin{aligned} \|\mathbb{S}_\sharp(f\sigma)\|_{L^{p,\infty}(w)}^p &\lesssim \kappa^p \mathfrak{M}_{p,\text{weak}}^p \|f\|_{L^p(\sigma)} + \sum_{Q \in \mathcal{Q}_k} w(Q)^{1-p} (\mathfrak{T}_p \|f1_Q\|_{L^p(\sigma)} w(Q)^{1/p'})^p \\ &= \kappa^p \mathfrak{M}_{p,\text{weak}}^p \|f\|_{L^p(\sigma)} + \mathfrak{T}_p^p \sum_{Q \in \mathcal{Q}_k} \|f1_Q\|_{L^p(\sigma)}^p \\ &\leq (\kappa^p \mathfrak{M}_{p,\text{weak}}^p + \mathfrak{T}_p^p) \|f\|_{L^p(\sigma)}, \end{aligned}$$

by the disjointness of the $Q \in \mathcal{Q}_k$ in the last step.

7. FIRST STEPS IN THE PROOF OF THE STRONG TYPE INEQUALITY

We start preparing for the proof of

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)} \lesssim (\kappa\mathfrak{M}_p + \mathfrak{T}_p + \mathfrak{N}_p)\|f\|_{L^p(\sigma)}.$$

In this section, we make an estimate of the form

$$\|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)} \lesssim \kappa\mathfrak{M}_p\|f\|_{L^p(\sigma)} + \delta\|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)} + \text{remainder},$$

where the second term with the small parameter δ may be absorbed to the right, and the ‘remainder’ will be controlled in terms on $(\mathfrak{T}_p + \mathfrak{N}_p)\|f\|_{L^p(\sigma)}$ in the following section, which contains the core of the argument.

We begin with

$$\begin{aligned} \|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)}^p &= \sum_{k \in \mathbb{Z}} \int_{\Omega_{k+1} \setminus \Omega_{k+2}} \mathbb{S}_\sharp(f\sigma)^p w \\ &\leq 4^p \sum_{k \in \mathbb{Z}} 2^{kp} w(\Omega_{k+1} \setminus \Omega_{k+2}) \\ &= 4^p \sum_{k \in \mathbb{Z}} 2^{kp} \sum_{Q \in \mathcal{Q}_k} w(Q \cap (\Omega_{k+1} \setminus \Omega_{k+2})) \\ &=: 4^p \sum_{k \in \mathbb{Z}} 2^{kp} \sum_{Q \in \mathcal{Q}_k} w(F_k(Q)). \end{aligned}$$

Note that the sets $F_k(Q)$ are pairwise disjoint. We first employ a similar reduction as in the weak-type case:

$$\begin{aligned} w(F_k(Q)) &\leq w(F_k(Q) \cap \{M(f\sigma) > \eta 2^k\}) + w(F_k(Q) \cap \{M(f\sigma) \leq \eta 2^k\}) \\ &=: w(F_k(Q) \cap \{M(f\sigma) > \eta 2^k\}) + w(E_k(Q)). \end{aligned}$$

Then

$$\begin{aligned} \sum_k 2^{kp} \sum_{Q \in \mathcal{Q}_k} w(F_k(Q) \cap \{M(f\sigma) > \eta 2^k\}) &\leq \eta^{-p} \sum_k \sum_{Q \in \mathcal{Q}_k} \int_{F_k(Q)} M(f\sigma)^p w \\ &\leq \eta^{-p} \|M(f\sigma)\|_{L^p(w)}^p \leq \eta^{-p} \mathfrak{M}_p^p \|f\|_{L^p(\sigma)}^p. \end{aligned}$$

We are left with

$$\begin{aligned} \sum_{k \in \mathbb{Z}} 2^{kp} \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) &\leq \sum_{k \in \mathbb{Z}} 2^{kp} \sum_{\substack{Q \in \mathcal{Q}_k \\ w(E_k(Q)) < \delta w(Q)}} \delta w(Q) \\ &\quad + \sum_{k \in \mathbb{Z}} 2^{kp} \sum_{\substack{Q \in \mathcal{Q}_k \\ w(E_k(Q)) \geq \delta w(Q)}} w(E_k(Q)), \end{aligned}$$

and the first sum on the right is immediately dominated by

$$\delta \sum_{k \in \mathbb{Z}} 2^{kp} w(\mathbb{S}_\sharp(f\sigma) > 2^k) \lesssim \delta \|\mathbb{S}_\sharp(f\sigma)\|_{L^p(w)}^p,$$

so this term can be absorbed after a suitable small choice of $\delta > 0$ depending only on p .

Now consider one of the remaining sets $E_k(Q)$ for $Q \in \mathcal{Q}_k$ and $w(E_k(Q)) \geq \delta w(Q)$. By the maximum principle, we can choose a linearization $\mathbb{L} = e^{i\vartheta(x)} \mathbb{S}_{\epsilon(x), v(x)}$ of \mathbb{S}_{\natural} with $v(x) \leq \ell(Q)$, so that, for $x \in E_k(Q)$,

$$\begin{aligned} 4 \cdot 2^k < \mathbb{S}_{\natural}(f\sigma)(x) &\leq 2\mathbb{L}(f\sigma)(x) + 2^k + (\zeta + n + 1)M(f\sigma)(x), \\ &\leq 2\mathbb{L}(f\sigma)(x) + 2^k + (\zeta + n + 1)\eta 2^k \\ &\leq 2\mathbb{L}(f\sigma)(x) + 2 \cdot 2^k, \end{aligned}$$

by choosing

$$\eta := (\zeta + n + 1)^{-1} = (2n + 2)^{-1} \gtrsim \kappa^{-1}.$$

Hence

$$\mathbb{L}(f\sigma)(x) \geq 2^k \quad \text{on } E_k(Q).$$

Notice that, by the disjointness of the sets $E_k(Q) \subset F_k(Q)$, we can globally define one linearization \mathbb{L} , which fulfills this condition on all $E_k(Q)$.

Thus, for $w(E_k(Q)) \geq \delta w(Q)$, we have

$$2^{kp} \leq \left(\frac{1}{w(E_k(Q))} \int_{E_k(Q)} \mathbb{L}(f\sigma)w \right)^p \leq \delta^{-p} \left(\frac{1}{w(Q)} \int_{E_k(Q)} \mathbb{L}(f\sigma)w \right)^p,$$

where

$$\mathbb{L} = e^{i\vartheta(x)} \mathbb{S}_{\epsilon(x), v(x)}, \quad v(x) \leq \ell(Q) \text{ on } E_k(Q).$$

We have proven that (absorbing the term $\delta \|\mathbb{S}_{\natural}(f\sigma)\|_{L^p(\sigma)}^p$, and using $\eta^{-1} \lesssim \kappa$)

$$\|\mathbb{S}_{\natural}(f\sigma)\|_{L^p(\sigma)}^p \lesssim \kappa^p \mathfrak{M}_p^p \|f\|_{L^p(\sigma)}^p + \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \left(\frac{1}{w(Q)} \int_{E_k(Q)} \mathbb{L}(f\sigma)w \right)^p.$$

(The dependence on δ has been neglected, since this is in any case a function of p only.)

8. STRONG-TYPE ESTIMATES: THE CORE

We are left to prove that

$$(8.1) \quad \sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \left| \frac{1}{w(Q)} \int_{E_k(Q)} \mathbb{L}(f\sigma)w \right|^p \lesssim (\mathfrak{T}_p + \mathfrak{N}_p)^p \|f\|_{L^p(\sigma)}.$$

We can make the additional assumption that all k in this sum are of the same parity; after all, there are just two such sums. By monotone convergence, we may also assume that all appearing cubes are contained in some maximal dyadic cube \overline{Q} . This allows to make the following construction:

Definition 8.2 (Principal cubes). We form the collection \mathcal{G} of *principal cubes* as follows. We let $\mathcal{G}_0 := \{\overline{Q}\}$ (the maximal cube that we consider), inductively

$$\mathcal{G}_k := \bigcup_{G \in \mathcal{G}_{k-1}} \{G' \subset G : \mathbb{E}_{G'}^\sigma |f| > 4\mathbb{E}_G^\sigma |f|, G' \text{ is a maximal such cube}\},$$

and then $\mathcal{G} := \bigcup_{k=0}^{\infty} \mathcal{G}_k$. For any dyadic Q ($\subset \overline{Q}$), we let

$$\Gamma(Q) := \text{the minimal principal cube containing } Q.$$

From the definition it follows that

$$\mathbb{E}_Q^\sigma |f| \leq 4\mathbb{E}_{\Gamma(Q)}^\sigma |f|.$$

We begin the analysis of (8.1). Recall that on $E_k(Q)$, we have $\mathbb{L} = \mathbb{L}(1_{Q \cdot})$. Thus we may dualize and split the cube Q to the result that

$$\begin{aligned} (8.3) \quad \int_{E_k(Q)} \mathbb{L}(f\sigma) w &= \int_Q \mathbb{L}^*(1_{E_k(Q)} w) f\sigma \\ &= \int_{Q \setminus \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f\sigma + \int_{Q \cap \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f\sigma. \end{aligned}$$

8.1. The part on $Q \setminus \Omega_{k+2}$. The first term is easy to estimate:

$$\begin{aligned} \left| \int_{Q \setminus \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f\sigma \right| &\leq \|1_Q \mathbb{L}^*(1_{E_k(Q)} w)\|_{L^{p'}(\sigma)} \|1_{Q \setminus \Omega_{k+2}} f\|_{L^p(\sigma)} \\ &\leq \mathfrak{T}_p w(Q)^{1/p'} \|1_{Q \setminus \Omega_{k+2}} f\|_{L^p(\sigma)}, \end{aligned}$$

and then

$$\begin{aligned} \sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \left| \frac{1}{w(Q)} \int_{Q \setminus \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f\sigma \right|^p \\ &\leq \mathfrak{T}_p^p \sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \frac{1}{w(Q)} \|1_{Q \setminus \Omega_{k+2}} f\|_{L^p(\sigma)}^p \\ &\leq \mathfrak{T}_p^p \sum_k \sum_{Q \in \mathcal{Q}_k} \|1_{Q \setminus \Omega_{k+2}} f\|_{L^p(\sigma)}^p \\ &\leq \mathfrak{T}_p^p \sum_k \|1_{\Omega_k \setminus \Omega_{k+2}} f\|_{L^p(\sigma)}^p = \mathfrak{T}_p^p \|f\|_{L^p(\sigma)}^p, \end{aligned}$$

recalling in the last step that the k sum is over either odd or even k only.

8.2. The part on $Q \cap \Omega_{k+2}$. We are left to estimate the integrals over $Q \cap \Omega_{k+2}$ as in the second term on the right of (8.3). Using $Q \in \mathcal{Q}_k$ and $\Omega_{k+2} = \bigcup_{R \in \mathcal{Q}_{k+2}} R$, as well as the nestedness of the collections \mathcal{Q}_k , we have

$$\int_{Q \cap \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f\sigma = \sum_{\substack{R \in \mathcal{Q}_{k+2} \\ R \subset Q}} \int_R \mathbb{L}^*(1_{E_k(Q)} w) f\sigma.$$

Now $E_k(Q) \subset \Omega_{k+2}^c \subset (R^{(\zeta)})^c$, where $\zeta = n+1$, so $\mathbb{L}^*(1_{E_k(Q)} w)$ is in fact constant on $R \in \mathcal{Q}_{k+2}$ (by the ‘smoothness property’ formulated in Lemma 3.5); thus

$$\int_R \mathbb{L}^*(1_{E_k(Q)} w) f\sigma = \int_R \mathbb{L}^*(1_{E_k(Q)} w) \sigma \cdot \mathbb{E}_R^\sigma f.$$

Splitting into the cases according to the size of $\mathbb{E}_R^\sigma |f|$ relative to $\mathbb{E}_{\Gamma(Q)}^\sigma |f|$, we can thus estimate

$$\begin{aligned}
 \left| \int_{Q \cap \Omega_{k+2}} \mathbb{L}^*(1_{E_k(Q)} w) f \sigma \right| &\leq 16 \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| \leq 16 \mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \cdot \mathbb{E}_{\Gamma(Q)}^\sigma |f| \\
 (8.4) \quad &+ \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16 \mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \cdot \mathbb{E}_R^\sigma |f|.
 \end{aligned}$$

8.3. The part with $\mathbb{E}_R^\sigma |f| \leq 16 \mathbb{E}_{\Gamma(Q)}^\sigma |f|$. We estimate the first sum on the right of (8.4). As a first step, by the disjointness of the $R \in \mathcal{Q}_{k+2}$ we have

$$\sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| \leq 16 \mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \leq \int_Q |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma.$$

Next, we make a manipulation involving random signs $\varepsilon_{k',Q'}$ indexed by pairs $k' \in \mathbb{Z}'$ and $Q' \in \mathcal{Q}_{k'}$. At almost every x ,

$$\begin{aligned}
 |\mathbb{L}^*(1_{E_k(Q)} 1_Q w)(x)| &\leq \mathbb{E}_\varepsilon \left| \sum_{k'} \sum_{Q' \in \mathcal{Q}_{k'}} \varepsilon_{k',Q'} \mathbb{L}^*(1_{E_{k'}(Q')} 1_Q w)(x) \right| \\
 &\leq \mathbb{E}_\varepsilon \left| \mathbb{L}^* \left(\sum_{k'} \sum_{Q' \in \mathcal{Q}_{k'}} \varepsilon_{k',Q'} 1_{E_{k'}(Q')} 1_Q w \right)(x) \right| \\
 &=: \mathbb{E}_\varepsilon |\mathbb{L}^*(\psi_\varepsilon 1_Q w)(x)|
 \end{aligned}$$

where, by the disjointness of the sets $E_{k'}(Q')$, we have $\|\psi_\varepsilon\|_\infty \leq 1$ for any choice of the signs $\varepsilon_{k',Q'}$.

We can then compute

$$\begin{aligned}
 &\sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \left(\frac{1}{w(Q)} \int_Q |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \cdot \mathbb{E}_{\Gamma(Q)}^\sigma |f| \right)^p \\
 &\leq \sum_{G \in \mathcal{G}} (\mathbb{E}_G^\sigma |f|)^p \sum_k \sum_{\substack{Q \in \mathcal{Q}_k \\ \Gamma(Q)=G}} w(E_k(Q)) \left(\frac{1}{w(Q)} \mathbb{E}_\varepsilon \int_Q |\mathbb{L}^*(\psi_\varepsilon 1_Q w)| \sigma \right)^p \\
 &\quad \quad \quad \text{(by the preceding considerations)} \\
 &\leq \sum_{G \in \mathcal{G}} (\mathbb{E}_G^\sigma |f|)^p \mathbb{E}_\varepsilon \sum_k \sum_{\substack{Q \in \mathcal{Q}_k \\ \Gamma(Q)=G}} w(E_k(Q)) \left(\frac{1}{w(Q)} \int_Q |\mathbb{L}^*(\psi_\varepsilon 1_Q w)| \sigma \right)^p \\
 &\quad \quad \quad \text{(by Hölder's inequality and the linearity of } \mathbb{E}_\varepsilon \text{)}
 \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{G \in \mathcal{G}} (\mathbb{E}_G^\sigma |f|)^p \mathbb{E}_\varepsilon \int_G \left(\sup_{Q \subset G} \frac{1_Q}{w(Q)} \int_Q |\mathbb{L}^*(\psi_\varepsilon 1_Q w)| \sigma \right)^p w \\
&\quad \text{(by the disjointness of the sets } E_k(Q) \text{)} \\
&\leq \sum_{G \in \mathcal{G}} (\mathbb{E}_G^\sigma |f|)^p \mathfrak{N}_p^p \sigma(G) \\
&\quad \text{(by the non-standard testing condition)} \\
&\lesssim \mathfrak{N}_p^p \|M_\sigma f\|_{L^p(\sigma)}^p \lesssim \mathfrak{N}_p^p \|f\|_{L^p(\sigma)}^p \\
&\quad \text{(by definition of principal cubes and universal maximal inequality.)}
\end{aligned}$$

8.4. The part with $\mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|$. We estimate the second term on the right of (8.4). Inserting $\sigma(R)^{\pm 1/p}$ into this sum and using Hölder's inequality, we have

$$\begin{aligned}
&\sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \cdot \mathbb{E}_R^\sigma |f| \\
&\leq \left(\sum_{\substack{R \in \mathcal{Q}_{k+2} \\ R \subset Q}} \sigma(R)^{-p'/p} \left[\int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \right]^{p'} \right)^{1/p'} \left(\sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \sigma(R) [\mathbb{E}_R^\sigma |f|]^p \right)^{1/p},
\end{aligned}$$

where the first factor is further dominated by

$$\left(\sum_{\substack{R \in \mathcal{Q}_{k+2} \\ R \subset Q}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)|^{p'} \sigma \right)^{1/p'} \leq \left(\int_Q |\mathbb{L}^*(1_{E_k(Q)} w)|^{p'} \sigma \right)^{1/p'} \leq \mathfrak{T}_p w(Q)^{1/p'}.$$

So altogether, and appealing to the testing condition (4.5),

$$\begin{aligned}
&\sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \left(\frac{1}{w(Q)} \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \int_R |\mathbb{L}^*(1_{E_k(Q)} w)| \sigma \cdot \mathbb{E}_R^\sigma |f| \right)^p \\
(8.5) \quad &\leq \mathfrak{T}_p^p \sum_k \sum_{Q \in \mathcal{Q}_k} w(E_k(Q)) \frac{1}{w(Q)} \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \sigma(R) [\mathbb{E}_R^\sigma |f|]^p \\
&\leq \mathfrak{T}_p^p \sum_k \sum_{Q \in \mathcal{Q}_k} \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \sigma(R) [\mathbb{E}_R^\sigma |f|]^p
\end{aligned}$$

The proof is completed by the following lemma:

Lemma 8.6. *Any given cube R appears at most once in the sum on the right of (8.5). For any two cubes $R_1 \subsetneq R_2$ appearing in this sum, we have $\mathbb{E}_{R_1}^\sigma |f| > 4\mathbb{E}_{R_2}^\sigma |f|$.*

Proof. We can prove both claims with one strike as follows. Let $R_1 \subset R_2$ be cubes appearing in (8.5), possibly equal. Thus for some cubes Q_i , we have

$$R_i \in \mathcal{Q}_{k_i+2}, \quad R_i \subset Q_i \in \mathcal{Q}_{k_i}, \quad \mathbb{E}_{R_i}^\sigma |f| > 16\mathbb{E}_{\Gamma(Q_i)}^\sigma |f|.$$

(In the equal case, we want to prove that $(k_1, Q_1) = (k_2, Q_2)$; in the unequal case, the estimate between the averages.) Note that if $k_1 = k_2$, then also $R_1 = R_2$ and $Q_1 = Q_2$, so there is nothing to prove. So let $k_1 > k_2$, thus by the restriction on the k -sum, in fact $k_1 \geq k_2 + 2$. Since the cubes $Q_1 \in \mathcal{Q}_{k_1}$ and $R_2 \in \mathcal{Q}_{k_2+2}$ intersect (on R_1) and $k_1 \geq k_2 + 2$, the nestedness property implies that $Q_1 \subset R_2$, and hence $\Gamma(Q_1) \subset \Gamma(R_2)$. Thus

$$\mathbb{E}_{R_1}^\sigma |f| > 16\mathbb{E}_{\Gamma(Q_1)}^\sigma |f| \geq 16\mathbb{E}_{\Gamma(R_2)}^\sigma |f| \geq 4\mathbb{E}_{R_2}^\sigma |f|.$$

For $R_1 = R_2$, this gives a contradiction, showing that the same R cannot arise in the sum more than once. And for $R_1 \subsetneq R_2$, this is precisely the asserted estimate. \square

From the lemma it follows that

$$\begin{aligned} \mathfrak{T}_p^p \sum_k \sum_{Q \in \mathcal{Q}_k} \sum_{\substack{R \in \mathcal{Q}_{k+2}, R \subset Q \\ \mathbb{E}_R^\sigma |f| > 16\mathbb{E}_{\Gamma(Q)}^\sigma |f|}} \sigma(R) [\mathbb{E}_R^\sigma |f|]^p \\ \lesssim \mathfrak{T}_p^p \|M_\sigma f\|_{L^p(\sigma)}^p \lesssim \mathfrak{T}_p^p \|f\|_{L^p(\sigma)}^p, \end{aligned}$$

and the proof is complete.

9. UNWEIGHTED WEAK-TYPE $(1,1)$ INEQUALITIES

We are now finished with the two-weight theory, and we start anew from a different corner of our proof scheme, see Figure 1, the weak-type $(1,1)$ estimates for Haar shift operators. Again, we need bounds that are effective in the complexity. The following estimate was proved in [7, Proposition 5.1], with the additional observation concerning shifts with separated scales made in [11, Theorem 5.2].

Definition 9.1. We say that a shift \mathbb{S} of complexity κ has its *scales separated* if all nonzero components $\mathbb{S}_Q, \mathbb{S}_{Q'}$ have $\log_2 \ell(Q) \equiv \log_2 \ell(Q') \pmod{\kappa+1}$. We likewise say that a subset \mathcal{Q} of the dyadic grid has *scales separated* if $\log_2 \ell(Q) \equiv \log_2 \ell(Q') \pmod{\kappa+1}$ for all $Q, Q' \in \mathcal{Q}$.

Proposition 9.2. *An L^2 -bounded Haar shift operator \mathbb{S} of complexity κ maps $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$ with norm at most $C_d \kappa$. If \mathbb{S} has scales separated, then the norm is at most C_d .*

We need a strengthening of this Proposition for duals \mathbb{L}^* of the all linearizations \mathbb{L} of the maximal truncation \mathbb{S}_\natural of \mathbb{S} . This type of result is not a classical one, and to prove it, we use a simplified version of the argument of [16] used to prove Carleson's Theorem on Fourier series [3].

Theorem 9.3. *For an L^2 -bounded Haar shift operator \mathbb{S} of complexity κ , we have the following estimate, uniform in $\lambda > 0$, and compactly supported and bounded functions f on \mathbb{R}^d*

$$(9.4) \quad \lambda |\{\mathbb{L}^* f > \lambda\}| \lesssim \kappa \|f\|_1,$$

where the inequality holds uniformly in choice of the linearization \mathbb{L} of \mathbb{S}_\sharp . If \mathbb{S} has its scales separated, we have the complexity-free bound

$$(9.5) \quad \lambda |\{\mathbb{L}^* f > \lambda\}| \lesssim \|f\|_1.$$

9.1. Case $\int \mathbb{S}_Q f = 0$. We begin proving the weak type $(1, 1)$ inequality with the additional hypothesis that the components of our shift \mathbb{S} all satisfy $\int \mathbb{S}_Q f = 0$. Note that this covers all cases relevant to us, except for the dual paraproduct case $\mathbb{S}_Q f = |Q|^{-1} \langle f, h_Q \rangle \mathbf{1}_Q$, which will be taken up in the next subsection.

We start off with the *Tree lemma*, with this terminology and notation adapted from [16].

Lemma 9.6 (Tree lemma). *Suppose that \mathcal{Q} is a collection of dyadic cubes, all contained in some Q_0 . Then*

$$\mathbb{L}_\mathcal{Q} f(x) := \sum_{\substack{Q \in \mathcal{Q} \\ \ell(Q) \leq v(x)}} \mathbb{S}_Q f(x)$$

satisfies

$$(9.7) \quad |\langle \mathbb{L}_\mathcal{Q} f, g \rangle| \lesssim \kappa \cdot \text{size}_f(\mathcal{Q}) \cdot \text{dense}_g(\mathcal{Q}) \cdot |Q_0|,$$

where

$$\text{size}_f(\mathcal{Q}) := \sup_{Q \in \mathcal{Q}} \left(\frac{1}{|Q|} \int_Q |f|^2 \right)^{1/2}, \quad \text{dense}_g(\mathcal{Q}) := \sup_{Q \in \mathcal{Q}} \sup_{Q' \supset Q} \frac{1}{|Q'|} \int_{Q'} |g|.$$

The factor κ may be omitted in (9.7) if the scales of \mathbb{S} are separated.

Remark 9.8. The notation of ‘size’ and ‘density’ are derived from [16]. But size in this context is simpler, related to the maximal function $(Mf^2)^{1/2}$. The double supremum in the definition of density is essential for the inequality (9.9) below.

Proof. We may assume that $Q_0 \in \mathcal{Q}$. For if not, let Q_i be the maximal cubes in \mathcal{Q} , all contained in Q_0 , and $\mathcal{Q}_i := \{Q \in \mathcal{Q} : Q \subset Q_i\}$. Clearly the size and density of each \mathcal{Q}_i is dominated by the corresponding number for \mathcal{Q} . Then we just write $\mathbb{L}_\mathcal{Q} = \sum_i \mathbb{L}_{\mathcal{Q}_i}$, use the estimate for each \mathcal{Q}_i , and sum up $\sum_i |Q_i| \leq |Q_0|$ in the end.

Also assume by approximation that the collection \mathcal{Q} is finite. Let \mathcal{P} consist of the minimal dyadic cubes $P \subset Q_0$ such that $P^{(1)}$ contains some element of \mathcal{Q} . The cubes in \mathcal{P} then form a partition of $\bigcup \{Q : Q \in \mathcal{Q}\}$. Define the maximal operator $\mathcal{M}_\mathcal{P}$ by

$$\mathcal{M}_\mathcal{P} \phi(x) := \sum_{P \in \mathcal{P}} |g(x)| \mathbf{1}_{P \cap \{\ell(P^{(1)}) \leq v(x)\}} \sup_{Q \supset P} \mathbb{E}_Q |\phi|.$$

Note that if $P \in \mathcal{P}$, then $P^{(1)}$ contains a cube $Q \in \mathcal{Q}$ with $P \subset Q$, whence

$$\mathbb{E}_P [|g| \mathbf{1}_{\{\ell(P^{(1)}) \leq v(\cdot)\}}] \leq 2^d \mathbb{E}_{P^{(1)}} [|g| \mathbf{1}_{\{\ell(P^{(1)}) \leq v(\cdot)\}}] \leq 2^d \text{dense}_g(\mathcal{Q}).$$

From this we conclude a particular restricted strong-type inequality for $\mathcal{M}_{\mathcal{P}}$:

$$\begin{aligned}
\int_{Q_0} \mathcal{M}_{\mathcal{P}} \phi &\leq \sum_{P \in \mathcal{P}} |P| \cdot \mathbb{E}_P[|g| \mathbf{1}_{\{\ell(P^{(1)}) \leq v(\cdot)\}}] \sup_{Q \supset P} \mathbb{E}_Q |\phi| \\
&\leq 2^d \text{dense}_g(\mathcal{Q}) \sum_{P \in \mathcal{P}} |P| \sup_{Q \supset P} \mathbb{E}_Q |\phi| \\
&\leq 2^d \text{dense}_g(\mathcal{Q}) \int_{Q_0} \mathcal{M} \phi \\
(9.9) \quad &\leq 2^d \text{dense}_g(\mathcal{Q}) |Q_0|^{1/2} \|\mathcal{M} \phi\|_2 \lesssim \text{dense}_g(\mathcal{Q}) |Q_0|^{1/2} \|\phi\|_2.
\end{aligned}$$

Now we start to estimate the expression in (9.7). For $x \in P \in \mathcal{P}$, we have

$$\mathbb{L}_{\mathcal{Q}} f(x) = \sum_{\substack{Q \in \mathcal{Q}, Q \supset P \\ \ell(Q) \leq v(x)}} \mathbb{S}_Q f(x) = \mathbf{1}_{\{\ell(P) < v(x)\}} \sum_{\substack{Q \in \mathcal{Q}, Q \supset P \\ \ell(Q) \leq v(x)}} \mathbb{S}_Q f(x),$$

since the summation is empty unless $\ell(P) < v(x)$. Let $P_{v(x)} \supsetneq P$ be the unique dyadic cube with $\ell(P_{v(x)}) = v(x)$. Then

$$\mathbf{1}_{\{\ell(P) < v(x)\}} \sum_{\substack{Q \in \mathcal{Q}, Q \supset P \\ \ell(Q) \leq v(x)}} \mathbb{S}_Q f(x) = \mathbf{1}_{\{\ell(P) < v(x)\}} \left(\sum_{\substack{Q \in \mathcal{Q} \\ Q \supset P}} \mathbb{S}_Q f(x) - \sum_{\substack{Q \in \mathcal{Q} \\ Q \supset P_{v(x)}}} \mathbb{S}_Q f(x) \right).$$

For any dyadic R and $x \in R$, we have

$$(9.10) \quad \sum_{\substack{Q \in \mathcal{Q} \\ Q \supset R}} \mathbb{S}_Q f(x) = \mathbb{E}_R \left(\sum_{Q \in \mathcal{Q}} \mathbb{S}_Q f \right) + \sum_{\substack{Q \in \mathcal{Q} \\ R \subsetneq Q \subset R^{(\kappa)}}} \left(\mathbb{S}_Q f(x) - \mathbb{E}_R(\mathbb{S}_Q f) \right),$$

which follows from the facts that $\mathbb{E}_R(\mathbb{S}_Q f) = 0$ for $Q \subset R$, and $\mathbb{S}_Q f$ is constant on R for $Q \supsetneq R^{(\kappa)}$. And here

$$\sum_{R \in \{P, P_{v(x)}\}} \left| \mathbb{E}_R \left(\sum_{Q \in \mathcal{Q}} \mathbb{S}_Q f \right) \right| \leq 2 \sup_{R \supset P} \mathbb{E}_R |\mathbb{S}_Q f|, \quad \mathbb{S}_{\mathcal{Q}} f := \sum_{Q \in \mathcal{Q}} \mathbb{S}_Q f.$$

For the second sum in (9.10), observe that we have

$$|\mathbb{E}_R(\mathbb{S}_Q f)| \leq \|\mathbb{S}_Q f\|_{\infty} \leq \frac{1}{|Q|} \int_Q |f| \leq \left(\frac{1}{|Q|} \int_Q |f|^2 \right)^{1/2} \leq \text{size}_f(\mathcal{Q})$$

for each term, and there are at most 4κ terms altogether for both $R \in \{P, P_{v(x)}\}$. If the scales of \mathbb{S} are separated, then there is at most one nonzero \mathbb{S}_Q with $R \subsetneq Q \subset R^{(\kappa)}$, so there are at most 4 nonzero terms with the mentioned estimate, instead of 4κ . This is the only place where the factor κ enters into the argument, and the rest of the proof can be modified to the case of separated scales by simply substituting 1 in place of κ .

Substituting back, we have shown that

$$|\mathbb{L}_{\mathcal{Q}} f(x)| \leq \mathbf{1}_{\{\ell(P) < v(x)\}} (2 \sup_{R \supset P} \mathbb{E}_R |\mathbb{S}_Q f| + 4\kappa \text{size}_f(\mathcal{Q})) \mathbf{1}_{Q_0}, \quad \forall x \in P \in \mathcal{P},$$

and hence

$$\begin{aligned} |\mathbb{L}_{\mathcal{Q}}f(x) \cdot g(x)| &\leq \sum_{P \in \mathcal{P}} |g(x)| \mathbf{1}_{P \cap \{\ell(P^{(1)}) \leq \nu(x)\}} \sup_{R \supset P} \mathbb{E}_R (2|\mathbb{S}_{\mathcal{Q}}f| + 4\kappa \text{size}_f(\mathcal{Q}) \mathbf{1}_{Q_0}) \\ &= \mathcal{M}_{\mathcal{P}}\phi(x), \quad \phi := 2|\mathbb{S}_{\mathcal{Q}}f| + 4\kappa \text{size}_f(\mathcal{Q}) \mathbf{1}_{Q_0}. \end{aligned}$$

An application of (9.9) then gives

$$|\langle \mathbb{L}_{\mathcal{Q}}f, g \rangle| \lesssim \text{dense}_g(\mathcal{Q}) \cdot |Q_0|^{1/2} \cdot \|\phi\|_2,$$

and

$$\begin{aligned} \|\phi\|_2 &\lesssim \|\mathbb{S}_{\mathcal{Q}}(1_{Q_0}f)\|_2 + \kappa \text{size}_f(\mathcal{Q}) \|\mathbf{1}_{Q_0}\|_2 \\ &\lesssim \|1_{Q_0}f\|_2 + \kappa \text{size}_f(\mathcal{Q}) |Q_0|^{1/2} \leq (1 + \kappa) \text{size}_f(\mathcal{Q}) |Q_0|^{1/2}. \end{aligned}$$

This completes the proof, recalling that in the case of separated scales we can take 1 in place of κ above. \square

The next two Lemmas present decompositions of collections \mathcal{Q} relative to the two quantities of density and size.

Lemma 9.11. *Let \mathcal{Q} be an arbitrary collection of dyadic cubes, $g \in L^1$, and $\delta > 0$. Then $\mathcal{Q} = \mathcal{Q}' \cup \bigcup_j \mathcal{Q}_j$ where $\text{dense}_g(\mathcal{Q}') \leq \delta \|g\|_1$ and $\mathcal{Q}_j = \{Q \in \mathcal{Q} : Q \subset Q_j\}$, where*

$$\sum_j |Q_j| \leq \delta^{-1}.$$

Proof. Let $\mathcal{Q}' = \{Q \in \mathcal{Q} : \sup_{Q' \supset Q} |Q'|^{-1} \int_{Q'} |g| \leq \delta \|g\|_1\}$. Then, by definition, there holds $\text{dense}_g(\mathcal{Q}') \leq \delta \|g\|_1$. Let Q_j be the maximal elements of $\mathcal{Q} \setminus \mathcal{Q}'$, and \mathcal{Q}_j be defined as in the statement. If $Q \in \mathcal{Q} \setminus \mathcal{Q}'$, then $|R(Q)|^{-1} \int_{R(Q)} |g| > \delta \|g\|_1$ for some dyadic $R(Q) \supset Q$. Let \mathcal{A} denote the maximal elements of $\{R(Q) : Q \in \mathcal{Q} \setminus \mathcal{Q}'\}$. We have

$$\sum_j |Q_j| = \sum_{Q \in \mathcal{A}} \sum_{j: R(Q_j) \subset Q} |Q_j| \leq \sum_{Q \in \mathcal{A}} |Q| \leq \sum_{Q \in \mathcal{A}} \frac{1}{\delta \|g\|_1} \int_Q |g| \leq \frac{1}{\delta},$$

where we used that $Q_j \subset R(Q_j)$, the disjointness of the cubes Q_j and the disjointness of the cubes of \mathcal{A} . \square

Lemma 9.12. *Let \mathcal{Q} be an arbitrary collection of dyadic cubes, $f \in L^2$, and $\sigma > 0$. Then $\mathcal{Q} = \mathcal{Q}' \cup \bigcup_j \mathcal{Q}_j$ where $\text{size}_f(\mathcal{Q}') \leq \sigma \|f\|_2$ and $\mathcal{Q}_j = \{Q \in \mathcal{Q} : Q \subset Q_j\}$, where*

$$\sum_j |Q_j| \leq \sigma^{-2}.$$

Proof. Let Q_j be the maximal elements of \mathcal{Q} with $(|Q_j|^{-1} \int_{Q_j} |f|^2)^{1/2} > \sigma \|f\|_2$, if any, and \mathcal{Q}_j be defined as in the statement. Then clearly $\mathcal{Q}' := \mathcal{Q} \setminus \bigcup_j \mathcal{Q}_j$ satisfies $\text{size}_f(\mathcal{Q}') \leq \sigma \|f\|_2$, and

$$\sum_j |Q_j| \leq \sum_j \frac{1}{\sigma^2 \|f\|_2^2} \int_{Q_j} |f|^2 \leq \frac{1}{\sigma^2 \|f\|_2^2} \|f\|_2^2 = \frac{1}{\sigma^2},$$

since the Q_j are disjoint by maximality. \square

Inductive application of the previous two Lemmas leads to the following general result on decomposition of an arbitrary collection of cubes \mathcal{Q} .

Lemma 9.13. *Let \mathcal{Q} be an arbitrary collection of dyadic cubes. Let n_0 be such that*

$$\text{dense}_g(\mathcal{Q}) \leq 2^{2n_0} \|g\|_1, \quad \text{size}_f(\mathcal{Q}) \leq 2^{n_0} \|f\|_2.$$

We can write a disjoint union

$$\mathcal{Q} = \mathcal{Q}^{-\infty} \cup \bigcup_{k=-\infty}^{n_0} \bigcup_j \mathcal{Q}_j^k$$

where

- (i) $\text{dense}_g(\mathcal{Q}_j^k) \leq 2^{2k} \|g\|_1$,
- (ii) $\text{size}_f(\mathcal{Q}_j^k) \leq 2^k \|f\|_2$,
- (iii) all cubes in \mathcal{Q}_j^k are contained in one Q_j^k , and

$$\sum_j |Q_j^k| \leq 8 \cdot 2^{-2k},$$

- (iv) both g and f vanish almost everywhere on all $Q \in \mathcal{Q}^{-\infty}$.

Proof. Using the previous Lemmas, we first split $\mathcal{Q} = \mathcal{Q}' \cup \bigcup_j \mathcal{Q}'_j$ where $\text{dense}_g(\mathcal{Q}') \leq 2^{2(n_0-1)} \|g\|_1$, and the cubes of \mathcal{Q}'_j are contained in Q'_j with $\sum_j |Q'_j| \leq 2^{-2(n_0-1)}$. Next, $\mathcal{Q}' = \mathcal{Q}'' \cup \bigcup_j \mathcal{Q}''_j$ where $\text{size}_f(\mathcal{Q}'') \leq 2^{n_0-1} \|f\|_2$, and the cubes of \mathcal{Q}''_j are contained in Q''_j with $\sum_j |Q''_j| \leq 2^{-2(n_0-1)}$. We re-enumerate the collections \mathcal{Q}'_j and \mathcal{Q}''_j as $\mathcal{Q}_j^{n_0}$, similarly for the containing cubes which satisfy $\sum_j |Q_j^{n_0}| \leq 8 \cdot 2^{-2n_0}$. Since $\mathcal{Q}_j^{n_0} \subset \mathcal{Q}$, these have density and size as required, and $\mathcal{Q}'' \subset \mathcal{Q}'$ has $\text{dense}_g(\mathcal{Q}'') \leq 2^{2(n_0-1)}$ and $\text{size}_f(\mathcal{Q}'') \leq 2^{n_0-1}$ by construction. We may thus iterate with (\mathcal{Q}, n_0) replaced by $(\mathcal{Q}'', n_0 - 1)$. If some cube $Q \in \mathcal{Q}$ is not chosen to any \mathcal{Q}_j^k at any phase of the iteration, this means that both $\int_Q |g| = \int_Q |f|^2 = 0$; these cubes constitute the collection $\mathcal{Q}^{-\infty}$. \square

Now we are prepared to prove the weak type $(1, 1)$ inequality for \mathbb{L}^* .

Proof of (9.4). We may assume that $\lambda = 1$ and that $|E| < \infty$, where $E := \{|\mathbb{L}^* g| > 1\}$. And consider the set $G = E \cap \{Mg \leq C_d \|g\|_1 / |E|\}$. Fixing C_d large enough (depending on the dimension d only), there holds

$$\frac{1}{2} |E| \leq |G| \leq \int_G |\mathbb{L}^* g| = \langle f, \mathbb{L}^* g \rangle = \langle \mathbb{L}_Q f, g \rangle,$$

where $|f| = \mathbf{1}_G$, and $\mathcal{Q} = \{Q \in \mathcal{D} : Q \cap G \neq \emptyset\}$; note that $\mathbb{S}_Q f = 0$ unless $Q \in \mathcal{Q}$.

The point of passing to the supplementary set G is that we have the estimate

$$\text{dense}_g(\mathcal{Q}) \leq \frac{C_d}{|E|} \|g\|_1.$$

And, as $|f| = 1_G$, it also follows that $\text{size}_f(\mathcal{Q}) \leq 1 = |G|^{-1/2} \|f\|_2$. We apply Lemma 9.13 to \mathcal{Q} , which yields the corresponding decomposition of $\mathbb{L}_{\mathcal{Q}}$. Note that $\langle \mathbb{L}_{\mathcal{Q} \sim \infty} f, g \rangle = 0$. Hence

$$\begin{aligned} |\{|\mathbb{L}^* g| > 1\}| &= |E| \lesssim |\langle \mathbb{L}_{\mathcal{Q}} f, g \rangle| \leq \sum_{k,j} |\langle \mathbb{L}_{\mathcal{Q}_j^k} f, g \rangle| \\ &\lesssim \kappa \sum_{k,j} \text{size}_j(\mathcal{Q}_j^k) \cdot \text{dense}_g(\mathcal{Q}_j^k) \cdot |Q_j^k| \\ &\lesssim \kappa \sum_k 2^k \|f\|_2 \cdot \min\{2^{2k}, |E|^{-1}\} \|g\|_1 \cdot \sum_j |Q_j^k| \\ &\lesssim \kappa \sum_k 2^k |E|^{1/2} \cdot \min\{2^{2k}, |E|^{-1}\} \|g\|_1 \cdot 2^{-2k} \lesssim \kappa \|g\|_1. \end{aligned}$$

If the scales of \mathbb{S} are separated, the factor κ does not appear in the application of the Tree Lemma 9.6, and we obtain instead that $|\{|\mathbb{L}^* g| > 1\}| \lesssim \|g\|_1$, as claimed. This completes the proof. \square

9.2. The case $\int \mathbb{S}_Q f \neq 0$. As mentioned, this only appears in the dual paraproduct case where $\mathbb{S}_Q f = |Q|^{-1} \langle f, h_Q \rangle \mathbf{1}_Q$, where h_Q is a Haar function on Q . In this case, the operator \mathbb{L} has the form

$$\mathbb{L}^* g(x) = \sum_{Q \in \mathcal{D}} |Q|^{-1} h_Q(x) \langle \mathbf{1}_{Q \cap \{\ell(Q) \leq v(Q)\}}, g \rangle = \sum_{Q \in \mathcal{D}} |Q|^{-1} h_Q(x) a_Q \langle \mathbf{1}_Q, |g| \rangle =: \tilde{\mathbb{S}}|g|(x),$$

where $|a_Q| \leq 1$ are some numbers, of course depending on g . However, the new shifts $\tilde{\mathbb{S}}f(x) := \sum_{Q \in \mathcal{D}} |Q|^{-1} h_Q(x) a_Q \langle \mathbf{1}_Q, f \rangle$ are uniformly bounded on L^2 with respect to the choice of the a_Q , hence by the weak-type estimate for untruncated shifts, also uniformly bounded from L^1 to $L^{1,\infty}$. Thus

$$\|\mathbb{L}^* g\|_{L^{1,\infty}} = \|\tilde{\mathbb{S}}g\|_{L^{1,\infty}} \lesssim \|g\|_{L^1},$$

and we are done in this case as well.

10. DISTRIBUTIONAL ESTIMATES

10.1. A John-Nirenberg Estimate. We recall this formulation from [11, Lemma 5.5]. Let \mathcal{D}_κ be a scales separated grid, as defined in Definition 9.1.

Definition 10.1. Let $\{\phi_Q : Q \in \mathcal{D}_\kappa\}$ be a collection of functions such that ϕ_Q is supported on Q and is constant on its \mathcal{D}_κ -children. For $R_0 \in \mathcal{D}_\kappa$ let $\phi_{R_0}^*$ be a maximal function

$$\phi_{R_0}^*(x) := \sup_{Q \in \mathcal{D}_\kappa : Q \ni x} \left| \sum_{\substack{R \in \mathcal{D}_\kappa \\ Q \subset R \subset R_0}} \phi_R(x) \right|.$$

Lemma 10.2. Let $\{\phi_Q : Q \in \mathcal{D}_\kappa\}$ be a collection of functions such that

- (1) ϕ_Q is supported on Q and constant on the \mathcal{D}_κ -children of Q ;
- (2) $\|\phi_Q\|_\infty \leq 1$;

(3) *There exists $\delta \in (0, 1)$ such that for all cubes $R \in \mathcal{D}_\kappa$*

$$\left| \left\{ x \in R : \phi_R^*(x) > 1 \right\} \right| \leq \delta |R|.$$

Then for all $R \in \mathcal{D}_\kappa$ and for all $t > 1$

$$\left| \left\{ x \in R : \phi_R^*(x) > t \right\} \right| \leq \delta^{(t-1)/2} |R|.$$

10.2. The Corona And Distributional Estimates. We need the important definition of the stopping cubes, and Corona Decomposition. The grid $\mathcal{D}_\kappa \subset \mathcal{D}$ has scales separated, as in Definition 9.1.

Definition 10.3. Let w be a weight and $Q \in \mathcal{D}_\kappa$. We set the w -stopping children $\mathcal{S}(Q)$ to be the maximal subcubes $Q' \subset Q$ with $Q' \in \mathcal{D}_\kappa$, so that

$$(10.4) \quad \frac{w(Q')}{|Q'|} \geq 4 \frac{w(Q)}{|Q|}.$$

Setting $\mathcal{S}_0 = \mathcal{S}(Q)$, and inductively setting $\mathcal{S}_{j+1} := \bigcup_{Q' \in \mathcal{S}_j} \mathcal{S}(Q')$, we refer to $\mathcal{S} := \bigcup_{j=0}^{\infty} \mathcal{S}_j$ as the w -stopping cubes of Q .

The w -Corona Decomposition of a collection of cubes \mathcal{Q} with $Q' \subset Q$ for all $Q' \in \mathcal{Q}$, consists of the w -stopping cubes \mathcal{S} , and collections of cubes $\{\mathcal{P}(S) : S \in \mathcal{S}\}$ so that (1) the collections $\mathcal{P}(S)$ form a disjoint decomposition of \mathcal{Q} , and (2) for all $S \in \mathcal{S}$, and $Q' \in \mathcal{P}(S)$, we have that S is the minimal element of \mathcal{S} that contains Q' . In particular, we have

$$\frac{w(Q')}{|Q'|} < 4 \frac{w(S)}{|S|}, \quad Q' \in \mathcal{P}(S), \quad S \in \mathcal{S}.$$

The previous definitions make sense for any weight w . Specializing to $w \in A_p$ leads to the following elementary, and essential, observations. The first is a familiar inequality, showing that an A_p weight cannot be too concentrated.

Lemma 10.5. *Let $w \in A_p$, Q is a cube and $E \subset Q$. We then have*

$$(10.6) \quad \left[\frac{|E|}{|Q|} \right]^p \|w\|_{A_p}^{-1} \leq \frac{w(E)}{w(Q)}.$$

Proof. The property that $w > 0$ a.e. allows us to write

$$\begin{aligned} \frac{|E|}{|Q|} &= \frac{\int_E w^{1/p}(x) w(x)^{-1/p} dx}{|Q|} \\ &\leq \frac{w(E)^{1/p} \sigma(Q)^{1/p'}}{|Q|} \\ &= \left[\frac{w(E)}{w(Q)} \right]^{1/p} \frac{w(Q)^{1/p} \sigma(Q)^{1/p'}}{|Q|} \end{aligned}$$

which proves the Lemma. \square

The second is a direct application of the previous assertion.

Lemma 10.7. *Let $w \in A_p$, Q a cube, and \mathcal{S} the w -stopping cubes for Q . Then, we have*

$$(10.8) \quad \sum_{S \in \mathcal{S}} w(S) \lesssim \|w\|_{A_p} w(Q).$$

Proof. It follows from (10.4) that we have that the union of the stopping children $\mathcal{S}(Q)$ has Lebesgue measure at most $\frac{1}{4}|Q|$. Applying (10.6) to the set $E = Q \setminus \bigcup_{S \in \mathcal{S}(Q)} S$, it follows that

$$\frac{w(E)}{w(Q)} \geq (3/4)^p \|w\|_{A_p}^{-1}.$$

Thus, we have

$$\sum_{S \in \mathcal{S}(Q)} w(S) \leq (1 - c\|w\|_{A_p}^{-1})w(Q).$$

Recalling the inductive definition of the w -stopping children, we see that our estimate holds. \square

10.3. The Distributional Estimates. We combine the John-Nirenberg and Corona Decomposition to obtain the crucial distributional estimates: The operator \mathbb{L} , decomposed using the Corona, satisfies exponential distributional inequalities. We will then strongly use the exponential moments to control certain L^p norms in the next section. Our estimates are two fold. The first inequalities involve the Lebesgue measure, which are the important intermediate step to obtain the second inequalities for the σ measure.

Definition 10.9. For $w \in A_p$, and integers $\alpha, \kappa \in \mathbb{N}$, we will say that a collection \mathcal{Q} of cubes is (w, α, κ) -adapted if these conditions hold. First, for any $Q, Q' \in \mathcal{Q}$, we have $\log_2 \ell(Q) = \log_2 \ell(Q') \pmod{\kappa + 1}$. Second, we have

$$(10.10) \quad 2^\alpha \leq \frac{w(Q)}{|Q|} \left[\frac{\sigma(Q)}{|Q|} \right]^{p-1} < 2^{\alpha+1}, \quad Q \in \mathcal{Q}.$$

We only need to consider $0 \leq \alpha \leq \lceil \log_2 \|w\|_{A_p} \rceil$.

Lemma 10.11. *Let $w \in A_p$, with dual measure σ , and let Q be a cube. For integers α, κ , let \mathcal{Q} be a collection of cubes contained in Q which are (σ, α, κ) -adapted. Construct the w -Corona Decompositions $\mathcal{P}(S)$, $S \in \mathcal{S}$ of \mathcal{Q} . Let \mathbb{S} be a Haar Shift operator of complexity κ , with $\|\mathbb{S}\|_{L^2 \rightarrow L^2} = 1$. We have these estimates, uniform in $t \geq 1$, choices of linearizations \mathbb{L} , functions φ with $\|\varphi\|_\infty \leq 1$, and $S \in \mathcal{S}$:*

$$(10.12) \quad \left| \left\{ x \in S : \mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_S \varphi w) \geq Kt \frac{w(S)}{|S|} \right\} \right| \lesssim 2^{-t} |S|,$$

$$(10.13) \quad \sigma \left(\left\{ x \in S : \mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_S \varphi w) \geq Kt \frac{w(S)}{|S|} \right\} \right) \lesssim 2^{-t} \sigma(S).$$

where K is a dimensional constant.

Observe that the condition that \mathcal{Q} be (σ, α, κ) -adapted implies in particular that all $\mathbb{L}_{\mathcal{Q}'}$, with $\mathcal{Q}' \subset \mathcal{Q}$, can be viewed as linearizations of shifts with separated scales; this will place the stronger conclusion (9.5) of Theorem 9.3 at our disposal, allowing us to get the stated complexity-free estimates (10.12) and (10.13). The proof given below follows almost to the letter the proof of [11, Lemma 5.6].

10.4. Proof of the distributional estimate for the Lebesgue measure. We begin with the bound (10.12). We aim to apply the John-Nirenberg estimate, Lemma 10.2. To this end, write

$$\begin{aligned} \mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_S \varphi w)(y) &= \sum_{Q \in \mathcal{P}(S)} \int_{Q \cap \{\varepsilon(x) \leq |Q| \leq v(x)\}} e^{i\vartheta(x)} s_Q(x, y) \mathbf{1}_S(x) \varphi(x) w(x) dx \\ &=: \sum_{Q \in \mathcal{P}(S)} \phi_Q(y), \end{aligned}$$

and

$$\phi_R^*(x) := \sup_{Q \in \mathcal{D}_\kappa: Q \ni x} \left| \sum_{Q' \in \mathcal{P}(S): Q \subsetneq Q' \subset R} \phi_{Q'}(x) \right|.$$

for $R \in \mathcal{D}_\kappa$ (note that $\phi_R^* \equiv 0$, if $R \not\subset S$). Then $|\mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_S \varphi w)| \leq \phi_S^*$, so that it suffices to prove (10.12) for ϕ_S^* instead of $\mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_S \varphi w)$. The condition (1) of Lemma 10.2 is clear: the function ϕ_Q is supported on Q , since the kernel s_Q is supported on $Q \times Q$. Also, ϕ_Q is constant on \mathcal{D}_κ -children of Q . The condition (2), that $\|\phi_Q\|_\infty \leq 1$, need not hold for cubes $Q \in \mathcal{P}(S)$, so we have to split these cubes into countably many subfamilies. For $\beta \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$, let $\mathcal{P}_\beta(S)$ consist of all cubes $Q \in \mathcal{P}(S)$ such that

$$4^{-\beta} \frac{w(S)}{|S|} \leq \frac{w(Q)}{|Q|} < 4^{-\beta+1} \frac{w(S)}{|S|}.$$

For every $Q \in \mathcal{P}(S)$ it holds that

$$\frac{w(Q)}{|Q|} < 4 \frac{w(S)}{|S|},$$

which means that every $Q \in \mathcal{P}(S)$ is contained in $\mathcal{P}_\beta(S)$ for some $\beta \in \mathbb{Z}_+$. Defining $\phi_{R,\beta}^*$ in the same manner as ϕ_R^* , only replacing $\mathcal{P}(S)$ by $\mathcal{P}_\beta(S)$ in the definition, we also have

$$\phi_S^* \leq \sum_{\beta \in \mathbb{Z}_+} \phi_{S,\beta}^*.$$

This will allow us to prove (10.12) for the functions $\phi_{S,\beta}^*$ separately. If $Q \in \mathcal{P}_\beta(S)$, we have

$$|\phi_Q(y)| \leq \frac{1}{|Q|} \int_{Q \cap \{\varepsilon(x) \leq |Q| \leq v(x)\}} w(x) dx \leq 4^{-\beta+1} \frac{w(S)}{|S|},$$

by definition of $\mathcal{P}_\beta(S)$. Thus condition (2) of Lemma 10.2 holds for the normalized functions $4^{\beta-1}|S|w(S)^{-1}\phi_Q$, $Q \in \mathcal{P}_\beta(S)$. In order to verify condition (3) for $\phi_{R,\beta}^*$, we need the weak-type (1,1) inequality for \mathbb{L}^* obtained above. To use this inequality, we first have to show that the set $\{x \in R : \phi_{R,\beta}^*(x) > \lambda\}$ is a subset of

$$\{x \in R : |\mathbb{L}_{\mathcal{Q}_\beta(R)}^*(\mathbf{1}_R \varphi w)(x)| > \lambda\}$$

for some subcollection $\mathcal{Q}_\beta(R) \subset \mathcal{P}_\beta(S)$. Let Q_1, Q_2, \dots be the maximal subcubes $Q \subset R$ in \mathcal{D}_k satisfying

$$\left| \sum_{Q' \in \mathcal{P}_\beta(S) : Q \subsetneq Q' \subset R} \phi_{Q'}(x) \right| > \lambda.$$

Note that every point in $\{x \in R : \phi_{R,\beta}^*(x) > \lambda\}$ is contained in exactly one Q_j and, in fact, $\{x \in R : \phi_{R,\beta}^*(x) > \lambda\} = \bigcup_j Q_j$. Then let

$$\mathcal{Q}_\beta(R) := \bigcup_{j=1}^{\infty} \{Q \in \mathcal{P}_\beta(S) : Q_j \subsetneq Q \subset R\}.$$

Now suppose $\phi_{R,\beta}^*(x) > \lambda$, and let Q_j be the unique cube containing x . Then, by definition,

$$\begin{aligned} \lambda &< \left| \sum_{Q \in \mathcal{P}_\beta(S) : Q_j \subsetneq Q \subset R} \phi_Q(x) \right| \\ &=: \left| \sum_{Q \in \mathcal{P}_\beta(S) : Q_j \subsetneq Q \subset R} \int_{Q \cap \{\varepsilon(y) \leq |Q| \leq v(y)\}} e^{i\vartheta(y)} s_Q(y, x) \mathbf{1}_S(y) \varphi(y) w(y) dy \right|. \end{aligned}$$

Since every cube in the sum is a subset of R , we may replace $\mathbf{1}_S$ by $\mathbf{1}_R$ above. The result is precisely $|\mathbb{L}_{\mathcal{Q}_\beta(R)}^*(\mathbf{1}_R \varphi w)(x)|$.

Let $K \geq 1$ be the dimensional constant of Theorem 9.3 for the weak-type inequality (9.5) for shifts with scales separated. Recalling that this separation is satisfied in our situation, we have in particular that

$$|\{\mathbb{L}_{\mathcal{Q}'}^* f > \lambda\}| \leq \frac{K \|f\|_1}{\lambda}, \quad \lambda > 0,$$

for any collection $\mathcal{Q}' \subset \mathcal{Q}$. Choosing $\lambda = 2 \cdot |S|^{-1}w(S)4^{-\beta+1}K$ now yields

$$\begin{aligned} |\{x \in R : \phi_{R,\beta}^*(x) > \lambda\}| &\leq |\{x \in R : |\mathbb{L}_{\mathcal{Q}_\beta(R)}^*(\mathbf{1}_R \varphi w)(x)| > \lambda\}| \\ &\leq \frac{K|S|\|\mathbf{1}_R w\|_1}{2 \cdot w(S)4^{-\beta+1}K} = \frac{|S| \cdot w(R)}{2 \cdot w(S)4^{-\beta+1}}. \end{aligned}$$

If $R \in \mathcal{P}_\beta(S)$, we immediately obtain $|\{x \in R : \phi_{R,\beta}^*(x) > \lambda\}| \leq |R|/2$. However, Lemma 10.2 requires the same estimate for all $R \in \mathcal{D}_\kappa$. Let R_1, R_2, \dots be the maximal cubes of $\mathcal{P}_\beta(S)$

inside R . Then $\phi_{R,\beta}^*$ is supported on the disjoint union of the cubes R_j , and $\phi_{R,\beta}^*(x) = \phi_{R_j,\beta}^*(x)$ for $x \in R_j$. Hence

$$|\{x \in R : \lambda^{-1}\phi_{R,\beta}^*(x) > 1\}| = \sum_{j=1}^{\infty} |\{x \in R_j : \phi_{R_j,\beta}^*(x) > \lambda\}| \leq \sum_{j=1}^{\infty} \frac{|R_j|}{2} \leq \frac{|R|}{2}.$$

Since $\lambda \geq |S|^{-1}w(S)4^{-\beta+1}$, we still have $\|\lambda^{-1}\phi_Q\|_{\infty} \leq 1$, so that finally the functions $\lambda^{-1}\phi_Q$, and the related $\lambda^{-1}\phi_{R,\beta}^*$, satisfy all requirements of Lemma 10.2. We conclude that

$$|\{x \in S : \phi_{S,\beta}^*(x) > t\lambda\}| \leq 2^{-(t-1)/2}|S|, \quad t > 1.$$

Note that the inequality is trivial for $0 < t < 1$, so that it holds for all $t > 0$. Recalling the definition of λ and rescaling t we can rewrite the previous as

$$(10.14) \quad \left| \left\{ x \in S : \phi_{S,\beta}^*(x) > 16t \frac{w(S)}{|S|} \right\} \right| \leq \sqrt{2} \cdot 2^{-t4^{\beta}/\kappa} |S|, \quad t > 0.$$

To finish the proof of (10.12), we use $\phi_S^* \leq \sum \phi_{S,\beta}^*$ to estimate

$$\begin{aligned} |S|^{-1} \left| \left\{ x \in S : \phi_S^*(x) > 16t \frac{w(S)}{|S|} \right\} \right| \\ \leq |S|^{-1} \sum_{\beta=0}^{\infty} \left| \left\{ x \in S : \phi_{S,\beta}^*(x) > 16 \cdot 2^{-\beta-1} t \frac{w(S)}{|S|} \right\} \right| \\ \leq \sum_{\beta=0}^{\infty} \sqrt{2} \cdot 2^{-t \cdot 2^{\beta-1}/\kappa} \stackrel{(*)}{\leq} \sqrt{2} \sum_{\beta=0}^{\infty} 2^{-t/2\kappa-\beta} = 2\sqrt{2} \cdot 2^{-t/2\kappa}. \end{aligned}$$

The inequality $(*)$ only holds in case $t \geq 2\kappa$, but for $t < 2\kappa$ the inequality just obtained is trivial. Replacing t by Kt proves (10.12).

10.5. Proof of the distributional estimate for the σ measure. In order to prove (10.13), we need the assumption that our cube family is (σ, α, κ) -adapted. Namely, recalling also the definition of $\mathcal{P}_{\beta}(S)$, we have the inequalities

$$2^{\alpha} 4^{\beta-1} \frac{|S|}{w(S)} \leq \left[\frac{\sigma(Q)}{|Q|} \right]^{p-1} \leq 2^{\alpha+1} 4^{\beta} \frac{|S|}{w(S)}, \quad Q \in \mathcal{P}_{\beta}(S).$$

Thus the measure σ is not impossibly far away from Lebesgue measure on the cubes $Q \in \mathcal{P}_{\beta}(S)$, and this allows us to make use of the previously established estimate for the Lebesgue measure. As was already noted during the proof of (10.12), any set of the form $\{x \in S : \phi_{S,\beta}^*(x) > \lambda\}$ can be expressed as the disjoint union of the *maximal* cubes $Q \in \mathcal{D}_{\kappa}$ for which $Q \subset S$ and

$$(10.15) \quad \left| \sum_{Q' \in \mathcal{P}_{\beta}(S) : Q \subset Q' \subset S} \phi_Q \right| > \lambda.$$

Apply this with $\lambda = 20t \cdot w(S)/|S|$, and let Q_1, Q_2, \dots be the maximal cubes mentioned above. Note that it cannot happen that $Q_j = S$, since then the summation condition in (10.15) would be empty ($S \subsetneq Q' \subset S$), and the estimate (10.15) could not possibly hold. Thus $Q_j \subsetneq S$ for every j , but there is no reason why $Q_j \in \mathcal{P}_\beta(S)$. Instead, $\tilde{Q}_j \in \mathcal{P}_\beta(S)$, where \tilde{Q}_j denotes the parent (in \mathcal{D}_k) of Q_j . This follows from the maximality of Q_j and the fact that the summation in (10.15) is only over $Q' \in \mathcal{P}_\beta(S)$. Now let

$$E_\beta(t) := \bigcup_{j=1}^{\infty} \tilde{Q}_j.$$

This union may be assumed disjoint, since it is anyway the disjoint union of the maximal cubes among the cubes $\tilde{Q}_j \in \mathcal{P}_\beta(S)$. Recalling the estimate $|\phi_Q(x)| \leq 4^{-\beta+1}w(S)|S|^{-1}$ used already in the proof of (10.12), we may now estimate

$$\begin{aligned} \left| \sum_{Q' \in \mathcal{P}_\beta(S): \tilde{Q}_j \subsetneq Q' \subset S} \phi_{Q'}(x) \right| &\geq \left| \sum_{Q' \in \mathcal{P}_\beta(S): Q_j \subsetneq Q' \subset S} \phi_{Q'}(x) \right| - |\phi_{\tilde{Q}_j}(x)| \\ &> 20t \frac{w(S)}{|S|} - 4 \cdot 4^{-\beta} \frac{w(S)}{|S|} \stackrel{(*)}{\geq} 16t \frac{w(S)}{|S|}, \quad x \in \tilde{Q}_j. \end{aligned}$$

Of course, $(*)$ only holds in case $t \geq 4^{-\beta}$. One should also observe that we needed here the fact that the sum in the first expression is constant on \tilde{Q}_j , the \mathcal{D}_κ child of the first cubes Q' in the actual summation: this becomes essential, if $x \in \tilde{Q}_j$ is not in Q_j to begin with, since we only have the inequality (10.15) for $x \in Q_j$. The previous estimate now shows that $E_\beta(t) \subset \{x : \phi_{S,\beta}(x) > 16t \cdot w(S)/|S|\}$ for $t \geq 4^{-\beta}$, whence

$$|E_\beta(t)| \leq \sqrt{2} \cdot 2^{-t4^\beta/\kappa} |S| \leq 2 \cdot 2^{-t4^\beta/\kappa} |S|$$

by (10.14). For $t < 4^{-\beta}$ this estimate says nothing, so the same bound holds for all $t > 0$. Now sum over the disjoint cubes that form $E_\beta(t)$ and use $\sigma(Q) \leq (2^{\alpha+1}4^\beta w(S)^{-1}|S|)^{1/(p-1)}|Q|$ to obtain

$$\begin{aligned} \sigma(E_\beta(t)) &\leq \left(2^{\alpha+1}4^\beta \frac{|S|}{w(S)}\right)^{1/(p-1)} |E_\beta(t)| \leq \left(2^{\alpha+1}4^\beta \frac{|S|}{w(S)}\right)^{1/(p-1)} 2 \cdot 2^{-t4^\beta/\kappa} |S| \\ &\leq (2 \cdot 4^\beta)^{1/(p-1)} \frac{\sigma(S)}{|S|} \cdot 2 \cdot 2^{-t4^\beta/\kappa} |S| = 2^{p'} \cdot 4^{\beta/(p-1)} \cdot 2^{-t4^\beta/\kappa} \sigma(S). \end{aligned}$$

In the last inequality we used again the fact that the cubes in $\mathcal{P}(S)$ are (σ, α, κ) -adapted. The estimate for $\sigma(\{x \in S : \phi_S^*(x) > 20t \cdot w(S)/|S|\})$ is now obtained in the same manner as

at the end of the proof of (10.12):

$$\begin{aligned}
& \sigma(S)^{-1} \left| \left\{ x \in S : \phi_S^*(x) > 20t \frac{w(S)}{|S|} \right\} \right| \\
& \leq \sigma(S)^{-1} \sum_{\beta=0}^{\infty} \left| \left\{ x \in S : \phi_{S,\beta}^*(x) > 20 \cdot 2^{-\beta-1} t \frac{w(S)}{|S|} \right\} \right| \\
(10.16) \quad & \leq 2^{p'} \sum_{\beta=0}^{\infty} 4^{\beta/(p-1)} 2^{-t2^{\beta-1}/2K} = 2^{p'} \sum_{\beta=0}^{\infty} 2^{2\beta/(p-1)-t2^{\beta}/2K}.
\end{aligned}$$

If $t < 2K$, we clearly have

$$\left| \left\{ x \in S : \phi_S^*(x) > 20t \frac{w(S)}{|S|} \right\} \right| \leq 2 \cdot 2^{-t/2K} \sigma(S).$$

We wish to prove a similar bound in case $t \geq 2K$ by estimating the last series on line (10.16). Since $2^\beta \geq (2/(p-1) + 1)\beta + 1$ for $\beta \geq \beta_p$, we have

$$2\beta/(p-1) - t2^\beta/2K \leq -\beta - t/2K, \quad \beta \geq \beta_p$$

Thus

$$2^{p'} \sum_{\beta=0}^{\infty} 2^{2\beta/(p-1)-t2^\beta/2K} \leq 2^{p'} \left(\sum_{\beta=0}^{\beta_p-1} 2^{2\beta/(p-1)} + \sum_{\beta=\beta_p}^{\infty} 2^{-\beta} \right) 2^{-t/2K} \leq C_p 2^{-t/2K}.$$

Hence

$$\left| \left\{ x \in S : \phi_S^*(x) > 20t \frac{w(S)}{|S|} \right\} \right| \leq C_p \cdot 2^{-t/2K} \sigma(S).$$

in each case. This completes the proof of (10.13), and thus of Lemma 10.11.

11. VERIFYING THE TESTING CONDITIONS

The main Theorems of §4 state that we can reduce the verification of the estimates of the main technical Theorem 2.10 to this Lemma.

Lemma 11.1. *For an L^2 -bounded Haar shift operator \mathbb{S} with complexity κ , a choice of $1 < p < \infty$, and $w \in A_p$, we have these estimates uniform over selection of dyadic cube Q and choice of linearization \mathbb{L} of the maximal truncations \mathbb{S}_\natural .*

$$(11.2) \quad \|\mathbf{1}_Q \mathbb{L}^*(w\varphi \mathbf{1}_Q)\|_{L^{p'}(\sigma)} \lesssim \kappa \|w\|_{A_p} w(Q)^{1/p'}, \quad$$

$$(11.3) \quad \left\| \sup_{R \subset Q} \mathbf{1}_R \frac{1}{w(R)} \int_R |\mathbb{L}^*(\mathbf{1}_R \varphi w)(y)| \sigma(dy) \right\|_{L^{p'}(w)} \lesssim \kappa \|w\|_{A_p}^{1/(p-1)} \sigma(Q)^{1/p}.$$

In these estimates, $p' = p/(p-1)$ is the conjugate index, $\sigma = w^{1-p'}$, and φ is a measurable function with $\|\varphi\|_\infty \leq 1$.

We shall return to the proof of the Lemma above, turning here to the proofs of the two main technical estimates in Theorem 2.10.

Proof of the weak-type estimate (2.11). We apply the inequality (4.4) from Theorem 4.3 giving sufficient conditions for the weak-type inequality in the general two-weight case, and then Buckley's bound for the weak-type maximal inequality constant $\mathfrak{M}_{p,\text{weak}}$ and the bound (11.2) for the 'backward' testing constant \mathfrak{T}_p in the one-weight case:

$$\begin{aligned} \|\mathbb{S}_\natural(f\sigma)\|_{L^{p,\infty}(w)} &\lesssim (\kappa\mathfrak{M}_{p,\text{weak}} + \mathfrak{T}_p) \|f\|_{L^p(\sigma)} \\ &\lesssim \left(\kappa\|w\|_{A_p}^{1/p} + \kappa\|w\|_{A_p} \right) \|f\|_{L^p(\sigma)} \lesssim \kappa\|w\|_{A_p} \|f\|_{L^p(\sigma)}, \end{aligned}$$

noting that the second term dominates since $\|w\|_{A_p} \geq 1$. \square

Proof of the strong-type estimate (2.12). We apply the sufficient conditions in the two weight setting of Theorem 4.7, and then the available estimates for the testing constants in the one-weight situation: Buckley's bound for the strong-type maximal inequality constant \mathfrak{M}_p , and the bounds (11.2) and (11.3) for the 'backward' and nonstandard testing constants \mathfrak{T}_p and \mathfrak{N}_p :

$$\begin{aligned} \|\mathbb{S}_\natural(f\sigma)\|_{L^p(w)} &\lesssim \{\kappa\mathfrak{M}_p + \mathfrak{T}_p + \mathfrak{N}_p\} \|f\|_{L^p(\sigma)} \\ &\lesssim \left\{ \kappa\|w\|_{A_p}^{1/(p-1)} + \kappa\|w\|_{A_p} + \kappa\|w\|_{A_p}^{1/(p-1)} \right\} \|f\|_{L^p(\sigma)} \\ &\lesssim \kappa\|w\|_{A_p}^{\max\{1,1/(p-1)\}} \|f\|_{L^p(\sigma)}. \end{aligned}$$

In this case either the first and third, or the second term dominates, depending on whether $p < 2$ or $p > 2$. \square

We turn to the proof of lemma 11.1.

Proof of (11.2). The data is fixed: of index p , weight $w \in A_p$, cube Q , function φ bounded by 1, and Haar Shift \mathbb{S} of complexity κ .

The definition of \mathbb{L}^* is as in (3.4), and is in particular a sum over all dyadic cubes P . Now, the sum in (3.4) can obviously be restricted to a sum over P that intersect Q . Moreover, the sum over P that contain Q can be controlled, using a straight forward appeal to the size conditions on the Haar Shift operators, and the definition of A_p .

Thus, in what follows, we need only consider cubes which are contained in Q . We split these cubes into (σ, α, κ) -adapted subcollections, in the sense of Definition 10.9. Namely, we form the $\kappa + 1$ subcollections according to the value of $\log_2 \ell(Q) \bmod (\kappa + 1)$. And each of them is further split according to the unique number α such that

$$2^\alpha \leq \frac{w(Q)}{|Q|} \left(\frac{\sigma(Q)}{|Q|} \right)^{p-1} < 2^{\alpha+1},$$

where $0 \leq \alpha \leq \lceil \log_2 \|w\|_{A_p} \rceil$ is an integer.

For any one of these (σ, α, κ) -adapted subcollections, say \mathcal{Q} , apply the Definition 10.3, to get w -stopping cubes \mathcal{S} , and corona decomposition $\{\mathcal{P}(S) : S \in \mathcal{S}\}$. Define functions $\mathbb{F}_S := \mathbb{L}_{\mathcal{P}(S)}^*(w\varphi \mathbf{1}_S)$. For these functions, we have the distributional estimate (10.13), and it

remains to show that

$$(11.4) \quad \left\| \sum_{S \in \mathcal{S}} \mathbb{F}_S \right\|_{L^{p'}(\sigma)}^{p'} \lesssim 2^{\alpha/p} \|w\|_{A_p}^{1/p'} w(Q)^{1/p'}.$$

To conclude (11.2), we sum this estimate over α to get the upper bound $\|w\|_{A_p} w(Q)^{1/p'}$, and then over the $\kappa + 1$ possible values of $\log_2 \ell(Q') \pmod{\kappa + 1}$ to get the estimate (11.2) as stated, with the factor κ .

The distributional estimates are exponential in nature, which permit a facile estimate of this norm. Define level sets of the functions \mathbb{F}_S by

$$\begin{aligned} F_{S,0} &:= \left\{ |F_S| < \mathsf{K} \frac{w(S)}{|S|} \right\}, \\ F_{S,n} &:= \left\{ \mathsf{K} n \frac{w(S)}{|S|} \leq |F_S| < \mathsf{K} (n+1) \frac{w(S)}{|S|} \right\}, \quad n \geq 1, \end{aligned}$$

where K is the dimensional constant in (10.13). And then we estimate, by a familiar trick,

$$\begin{aligned} \left| \sum_{S \in \mathcal{S}} \mathbb{F}_S \right|^{p'} &= \left| \sum_{n=0}^{\infty} \sum_{S \in \mathcal{S}} (1+n)^{-2/p+2/p} \mathbb{F}_S \mathbf{1}_{F_{S,n}} \right|^{p'} \\ &\leq \left(\frac{\pi^2}{2} \right)^{p'/p} \sum_{n=0}^{\infty} (1+n)^{2p'/p} \left| \sum_{S \in \mathcal{S}} \mathbb{F}_S \mathbf{1}_{F_{S,n}} \right|^{p'} \\ &\leq \left(\frac{\pi^2}{2} \right)^{p'/p} \mathsf{K}^{p'} \sum_{n=0}^{\infty} (1+n)^{2p'/p+p'} \left(\sum_{S \in \mathcal{S}} \frac{w(S)}{|S|} \mathbf{1}_{F_{S,n}} \right)^{p'} \\ &\leq \left(\frac{\pi^2}{2} \right)^{p'/p} \mathsf{K}^{p'} \sum_{n=0}^{\infty} (1+n)^{2p'/p+p'} \left(\frac{4}{3} \right)^{p'} \sum_{S \in \mathcal{S}} \left[\frac{w(S)}{|S|} \right]^{p'} \mathbf{1}_{F_{S,n}}. \end{aligned}$$

Here, we use Hölder's inequality in the first inequality, and in the fourth, we use the fact that the averages $\frac{w(S)}{|S|}$ increase geometrically along the stopping intervals. This is wasteful in the powers on n , which is of no consequence for us.

We then prove (11.4) as follows, absorbing the dimensional constant K into the constants implicit in the notation \lesssim . We have

$$\begin{aligned} \left\| \sum_{S \in \mathcal{S}} \mathbb{F}_S \right\|_{L^{p'}(\sigma)}^{p'} &\lesssim \sum_{n=0}^{\infty} (1+n)^{2p'/p+p'} \sum_{S \in \mathcal{S}} \left[\frac{w(S)}{|S|} \right]^{p'} \sigma(F_{S,n}) \\ &\lesssim \sum_{n=0}^{\infty} (1+n)^{2p'/p+p'} 2^{-n} \sum_{S \in \mathcal{S}} \left[\frac{w(S)}{|S|} \right]^{p'} \sigma(S) \\ &\lesssim \sum_{S \in \mathcal{S}} w(S) \left[\frac{w(S)}{|S|} \right]^{p'-1} \frac{\sigma(S)}{|S|} \\ (11.5) \quad &\lesssim 2^{(p'-1)\alpha} \sum_{S \in \mathcal{S}} w(S) \lesssim 2^{(p'-1)\alpha} \|w\|_{A_p} w(Q). \end{aligned}$$

To pass to the second line, we use the distributional estimate (10.13), giving us a sum in n that is trivially bounded. From the third to the fourth line, we use the condition (10.10) of Definition 10.9 and then use the estimate (10.8). \square

Proof of (11.3). Fix the relevant data for this condition, and take a dyadic cube $R \subset Q$. We take $\tilde{\mathbb{L}}_R^*$ to be as in (3.4), but with the sum restricted to cubes that contain R . Then,

$$\sup_{R \subset Q} \frac{1}{w(R)} \int_R |\tilde{\mathbb{L}}_R^*(w\varphi \mathbf{1}_R)| \sigma(dy) \leq \sup_{R \subset Q} \frac{\sigma(R)}{|R|} \leq M(\sigma \mathbf{1}_Q).$$

And this last term is controlled by Buckley's estimate for the Maximal Function. Thus, when we consider the expression

$$\frac{1}{w(R)} \int_R |\mathbb{L}^*(\mathbf{1}_R \varphi w)(y)| \sigma(dy)$$

we can in addition assume that all cubes contributing to \mathbb{L}^* are contained in R , namely we can replace \mathbb{L}^* by $\mathbb{L}_{\mathcal{R}(R)}^*$, where $\mathcal{R}(R)$ is an appropriate collection of cubes contained in R .

To make a relevant estimate of this integral, we can consider \mathcal{Q} a collection of cubes $Q' \subset Q$ that are (w, α, κ) -adapted, where $0 \leq \alpha \leq \lceil \log_2 \|w\|_{A_p} \rceil$. Then, we will show that

$$(11.6) \quad \left\| \sup_{R \subset Q} \mathbf{1}_R \frac{1}{w(R)} \int_R |\mathbb{L}_{\mathcal{R}(R)}^*(\mathbf{1}_R \varphi w)(y)| \sigma(dy) \right\|_{L^p(w)} \lesssim 2^{\alpha/(p-1)} \sigma(Q)^{1/p},$$

where the collections $\mathcal{R}(R)$ are those cubes $Q' \in \mathcal{Q}$ with $Q' \subset R$. Summing this estimate over $0 \leq \alpha \leq \lceil \log_2 \|w\|_{A_p} \rceil$ and the $\kappa + 1$ possible values of $\log_2 \ell(Q') \bmod (\kappa + 1)$ will prove (11.3).

We then apply Definition 10.3, letting $\mathcal{S}(R)$ be a collection of stopping cubes for $\mathcal{R}(R)$, with Corona Decomposition $\{\mathcal{P}(S) : S \in \mathcal{S}(R)\}$. Then, we have

$$\begin{aligned} \frac{1}{w(R)} \int_R |\mathbb{L}_{\mathcal{R}(R)}^*(\mathbf{1}_R \varphi w)(y)| \sigma(dy) &\leq \frac{1}{w(R)} \sum_{S \in \mathcal{S}(R)} \int_R |\mathbb{L}_{\mathcal{P}(S)}^*(\mathbf{1}_R \varphi w)(y)| \sigma(dy) \\ &\lesssim \kappa \frac{1}{w(R)} \sum_{S \in \mathcal{S}(R)} \frac{w(S)}{|S|} \sigma(S) \\ &\lesssim \kappa 2^{\alpha/(p-1)} \frac{1}{w(R)} \sum_{S \in \mathcal{S}(R)} w(S) \left[\frac{|S|}{w(S)} \right]^{1/(p-1)}. \end{aligned}$$

Here, we use the estimate (10.13). Recalling that κ is only a dimensional constant, we henceforth absorb it from the notation. And so, to prove (11.6), we should show that

$$(11.7) \quad \left\| \sup_{R \subset Q} \mathbf{1}_R \frac{1}{w(R)} \sum_{S \in \mathcal{S}(R)} w(S) \left[\frac{|S|}{w(S)} \right]^{1/(p-1)} \right\|_{L^p(w)} \lesssim \sigma(Q)^{1/p}.$$

We estimate

$$\begin{aligned}
\sum_{S \in \mathcal{S}(R)} w(S) \left[\frac{|S|}{w(S)} \right]^{1/(p-1)} &= \sum_{k=0}^{\infty} \sum_{S \in \mathcal{S}_k(R)} w(S) \left[\frac{|S|}{w(S)} \right]^{1/(p-1)} \\
&\leq \sum_{k=0}^{\infty} \sum_{S \in \mathcal{S}_k(R)} w(S) \left[4^{-k} \frac{|R|}{w(R)} \right]^{1/(p-1)} \\
&\leq \sum_{k=0}^{\infty} w(R) \left[4^{-k} \frac{|R|}{w(R)} \right]^{1/(p-1)} \\
&\lesssim w(R) \left[\frac{|R|}{w(R)} \right]^{1/(p-1)} = w(R) [\mathbb{E}_R^w(w^{-1} \mathbf{1}_Q)]^{1/(p-1)},
\end{aligned}$$

where we used the fact that the terms $\frac{|S|}{w(S)}$ decrease geometrically along the levels of the stopping cubes. Therefore, to see the estimate (11.7), we should estimate

$$\|M_w(w^{-1} \mathbf{1}_Q)^{1/(p-1)}\|_{L^p(w)} = \|M_w(w^{-1} \mathbf{1}_Q)\|_{L^{p'}(w)}^{1/(p-1)} \lesssim \|w^{-1} \mathbf{1}_Q\|_{L^{p'}(w)}^{1/(p-1)} = \sigma(Q)^{1/p}$$

by the universal maximal function estimate. \square

12. VARIATIONS ON THE MAIN THEOREM

In this final section, we present a couple of variations of the main result, Theorem 1.2, where we allow the appearance of different weight characteristics than just $\|w\|_{A_p}$ in the norm estimates in $L^p(w)$. First, as a direct consequence of Theorem 1.2, we deduce the following result, which was conjectured by Lerner and Ombrosi [19, Conjecture 1.3] for untruncated operators T :

Corollary 12.1. *For T an $L^2(\mathbb{R}^d)$ bounded Calderón-Zygmund Operator and $1 < q < p < \infty$,*

$$\|T_{\natural} f\|_{L^p(w)} \leq C_{T,p,q} \|w\|_{A_q} \|f\|_{L^p(w)}.$$

Proof. Choose some $r \in (p, \infty)$. Since $\|w\|_{A_r} \leq \|w\|_{A_q}$, the weak-type estimate of Theorem 1.2 shows that

$$\begin{aligned}
\|T_{\natural} f\|_{L^{q,\infty}(w)} &\leq C_{T,q} \|w\|_{A_q} \|f\|_{L^q(w)}, \\
\|T_{\natural} f\|_{L^{r,\infty}(w)} &\leq C_{T,r} \|w\|_{A_r} \|f\|_{L^r(w)} \leq C_{T,r} \|w\|_{A_q} \|f\|_{L^r(w)}.
\end{aligned}$$

It suffices to apply the Marcinkiewicz interpolation theorem to the sublinear operator T_{\natural} to deduce the asserted strong-type bound in $L^p(w)$. (For $p \geq 2$, we could also have used the strong-type estimate of Theorem 1.2 and $\|w\|_{A_p} \leq \|w\|_{A_q}$, without any interpolation.) \square

In order to describe the other variant of the main theorem, we consider the A_{∞} characteristic

$$(12.2) \quad \|w\|_{A_{\infty}} := \sup_Q \frac{1}{w(Q)} \int_Q M(w \mathbf{1}_Q),$$

which has been introduced (with a different notation) by Wilson [26], and more recently used by Lerner [17]. One can show (see [10]) that $\|w\|_{A_\infty} \leq c_d \|w\|_{A_p}$ for all $p \in (1, \infty)$, and therefore the following estimate is an improvement of Theorem 1.2:

Theorem 12.3. *For T an $L^2(\mathbb{R}^d)$ bounded Calderón-Zygmund Operator and $1 < p < \infty$,*

$$\begin{aligned} \|T_{\natural} f\|_{L^{p,\infty}(w)} &\leq C_{T,p} \|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} \|f\|_{L^p(w)}, \\ \|T_{\natural} f\|_{L^p(w)} &\leq C_{T,p} (\|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} + \|w\|_{A_p}^{1/(p-1)}) \|f\|_{L^p(w)}. \end{aligned}$$

Note that the strong-type bound is a strict improvement of the corresponding bound in Theorem 1.2 only for $2 < p < \infty$. For $1 < p \leq 2$, the term $\|w\|_{A_p}^{1/(p-1)}$ dominates, and this is exactly the same bound as in Theorem 1.2.

An earlier result of this type is implicitly contained in the work of Lerner [17, Section 5.5]:

$$\|T_{\natural} f\|_{L^p(w)} \leq C_{T,p} \|w\|_{A_p}^{1/2} \|w\|_{A_\infty}^{1/2} \|f\|_{L^p(w)}, \quad 3 \leq p < \infty.$$

Our estimate improves on this, as both terms $\|w\|_{A_\infty}^{1/(p-1)} \leq \|w\|_{A_\infty}^{1/2}$ (for $p-1 \geq 2$) and

$$\|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} = \|w\|_{A_p}^{1/2} \|w\|_{A_\infty}^{1/2} \left(\frac{\|w\|_{A_\infty}}{\|w\|_{A_p}} \right)^{1/2-1/p}$$

are smaller than $\|w\|_{A_p}^{1/2} \|w\|_{A_\infty}^{1/2}$ for $3 \leq p < \infty$.

On the other hand, Theorem 12.3 fails to reproduce the following bound for untruncated operators T , which was recently obtained in [10]:

$$\|Tf\|_{L^2(w)} \lesssim \|w\|_{A_2}^{1/2} (\|w\|_{A_\infty} + \|w^{-1}\|_{A_\infty})^{1/2} \|f\|_{L^2(w)}.$$

We do not know whether this bound remains true for T_{\natural} in place of T . More generally, although the optimal $L^p(w)$ bounds in terms of $\|w\|_{A_p}$ coincide for T and T_{\natural} , we do not know if this is the case when allowing the more complicated dependence on both A_p and A_∞ characteristics.

We now discuss the proof of Theorem 12.3. Just like the proof of Theorem 1.2, thanks to the Representation Theorem 2.5, this is a consequence of an analogous result for Haar shifts:

Theorem 12.4. *Let \mathbb{S} be a Haar shift operator with complexity κ , a paraproduct, or a dual paraproduct. For $1 < p < \infty$ and $w \in A_p$, we have the estimates*

$$\begin{aligned} \|\mathbb{S}_{\natural} f\|_{L^{p,\infty}(w)} &\lesssim \kappa \|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} \|f\|_{L^p(w)}, \\ \|\mathbb{S}_{\natural} f\|_{L^p(w)} &\lesssim \kappa (\|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'} + \|w\|_{A_p}^{1/(p-1)}) \|f\|_{L^p(w)}. \end{aligned}$$

Sketch of proof. We can exploit the same two-weight estimates as before,

$$\begin{aligned} \|\mathbb{S}_{\natural}(f\sigma)\|_{L^{p,\infty}(w)} &\lesssim \{\kappa \mathfrak{M}_{p,\text{weak}} + \mathfrak{T}_p\} \|f\|_{L^p(\sigma)}, \\ \|\mathbb{S}_{\natural}(f\sigma)\|_{L^p(w)} &\lesssim \{\kappa \mathfrak{M}_p + \mathfrak{T}_p + \mathfrak{N}_p\} \|f\|_{L^p(\sigma)}, \end{aligned}$$

and it only remains to see how the A_∞ norm can be incorporated into the estimates of the testing constants appearing here. Recall that we have

$$\mathfrak{M}_{p,\text{weak}} \lesssim \|w\|_{A_p}^{1/p}, \quad \mathfrak{M}_p \lesssim \|w\|_{A_p}^{1/(p-1)}, \quad \mathfrak{N}_p \lesssim \kappa \|w\|_{A_p}^{1/(p-1)},$$

which are of admissible size for Theorem 12.4. The improvement over the earlier bounds comes from a more careful estimation of \mathfrak{T}_p .

Recall that the estimate $\mathfrak{T}_p \lesssim \kappa \|w\|_{A_p}$ is proven in Lemma 11.1, Eq. (11.2), and that this proof is reduced to the estimate (11.4). The final step in the proof of this estimate, (11.5), is an application of Lemma 10.7, which says that the collection \mathcal{S} of w -stopping cubes for a cube Q satisfies

$$\sum_{S \in \mathcal{S}} w(S) \lesssim \|w\|_{A_p} w(Q).$$

It is this last bound where $\|w\|_{A_p}$ can be replaced by $\|w\|_{A_\infty}$, as observed in [10]:

$$\sum_{S \in \mathcal{S}} w(S) \lesssim \|w\|_{A_\infty} w(Q).$$

Using this bound in (11.5), we obtain

$$\mathfrak{T}_p \lesssim \kappa \|w\|_{A_p}^{1/p} \|w\|_{A_\infty}^{1/p'},$$

and thus the estimates as asserted in Theorem 12.4. \square

Remark 12.5. It is shown in [10] that there is the sharper bound $\mathfrak{M}_p \lesssim (\|w\|_{A_p} \|\sigma\|_{A_\infty})^{1/p}$, where $\sigma = w^{-1/(p-1)}$. However, this does not allow us improve on the above estimates, as our bound for \mathfrak{N}_p , namely $\|w\|_{A_p}^{1/(p-1)}$, is already as large as Buckley's classical bound for \mathfrak{M}_p .

Remark 12.6. Another, perhaps more commonly used definition of the A_∞ characteristic is the following quantity introduced by Hruščev [5]:

$$\sup_Q \left(\frac{1}{|Q|} \int_Q w \right) \exp \left(\frac{1}{|Q|} \int_Q \log w^{-1} \right).$$

The A_∞ characteristic as defined in (12.2) is always smaller (up to a dimensional constant), and it can be much smaller for some weights (see [10] for details), so Theorems 12.3 and 12.4 are sharper when stated in terms of $\|w\|_{A_\infty}$ as in (12.2).

REFERENCES

- [1] Pascal Auscher, Steve Hofmann, Camil Muscalu, Terence Tao, and Christoph Thiele, *Carleson measures, trees, extrapolation, and $T(b)$ theorems*, Publ. Mat. **46** (2002), no. 2, 257–325.
- [2] Stephen M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), no. 1, 253–272.
- [3] Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.
- [4] David Cruz-Uribe, José María Martell, and Carlos Pérez, *Sharp weighted estimates for approximating dyadic operators*, Electron. Res. Announc. Math. Sci. **17** (2010), 12–19.

- [5] Sergei V. Hruščev, *A description of weights satisfying the A_∞ condition of Muckenhoupt*, Proc. Amer. Math. Soc. **90** (1984), no. 2, 253–257.
- [6] Richard Hunt, Benjamin Muckenhoupt, and Richard Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. **176** (1973), 227–251.
- [7] Tuomas Hytönen, *The sharp weighted bound for general Calderon-Zygmund operators* (2010), available at <http://arxiv.org/abs/1007.4330>.
- [8] Tuomas Hytönen, Michael T. Lacey, Maria Carmen Reguera, and Armen Vagharshakyan, *Weak and Strong-type estimates for Haar Shift Operators: Sharp power on the A_p characteristic* (2009), available at <http://www.arxiv.org/abs/0911.0713>.
- [9] Tuomas Hytönen, Michael T. Lacey, Maria Carmen Reguera, Eric T. Sawyer, Ignacio Uriarte-Tuero, and Armen Vagharshakyan, *Weak and Strong type A_p Estimates for Caldern-Zygmund Operators* (2010), available at <http://www.arxiv.org/abs/1006.2530>.
- [10] T. Hytönen and C. Pérez, *Sharp weighted bounds involving A_∞* , Manuscript in preparation (2011).
- [11] T. Hytönen, Carlos Pérez, S. Treil, and A. Volberg, *Sharp weighted estimates of the dyadic shifts and A_2 conjecture*, ArXiv e-prints, available at <http://arxiv.org/abs/1010.0755>.
- [12] Michael T. Lacey, Stefanie Petermichl, and Maria Carmen Reguera, *Sharp A_2 inequality for Haar shift operators*, Math. Ann. **348** (2010), no. 1, 127–141.
- [13] Michael T. Lacey, Eric T. Sawyer, and Ignacio Uriarte-Tuero, *Two Weight Inequalities for Discrete Positive Operators*, Submitted (2009), available at <http://www.arxiv.org/abs/0911.3437>.
- [14] ———, *A characterization of two weight norm inequalities for maximal singular integrals with one doubling measure*, A&PDE, to appear (2008), available at <http://arxiv.org/abs/0805.0246>.
- [15] ———, *Two Weight Inequalities for Maximal Truncations of Dyadic Calderón-Zygmund Operators*, Submitted (2009), available at <http://www.arxiv.org/abs/0911.3920>.
- [16] Michael T. Lacey and Christoph Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370.
- [17] Andrei K. Lerner, *Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals* (2010), available at <http://arxiv.org/abs/1005.1422>.
- [18] ———, *On some weighted norm inequalities for Littlewood-Paley operators*, Illinois J. Math. **52** (2007), no. 2, 653–666.
- [19] Andrei K. Lerner and Sheldy Ombrosi, *An extrapolation theorem with applications to weighted estimates for singular integrals*, Preprint (2010).
- [20] F. Nazarov, S. Treil, and A. Volberg, *Accretive system Tb -theorems on nonhomogeneous spaces*, Duke Math. J. **113** (2002), no. 2, 259–312.
- [21] ———, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc. **12** (1999), no. 4, 909–928.
- [22] Stefanie Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic*, Amer. J. Math. **129** (2007), no. 5, 1355–1375.
- [23] Eric T. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. **75** (1982), no. 1, 1–11.
- [24] Eric Sawyer, *A two weight weak type inequality for fractional integrals*, Trans. Amer. Math. Soc. **281** (1984), no. 1, 339–345.
- [25] Eric T. Sawyer, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), no. 2, 533–545.
- [26] J. Michael Wilson, *Weighted inequalities for the dyadic square function without dyadic A_∞* , Duke Math. J. **55** (1987), no. 1, 19–50.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, GUSTAF HÄLLSTRÖMIN
KATU 2B, FI-00014 HELSINKI, FINLAND

E-mail address: tuomas.hytonen@helsinki.fi

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
E-mail address: lacey@math.gatech.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, GUSTAF HÄLLSTRÖMIN
KATU 2B, FI-00014 HELSINKI, FINLAND

E-mail address: henri.martikainen@helsinki.fi

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, GUSTAF HÄLLSTRÖMIN
KATU 2B, FI-00014 HELSINKI, FINLAND

E-mail address: tuomas.orponen@helsinki.fi

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
E-mail address: mreguera@math.gatech.edu

DEPARTMENT OF MATHEMATICS & STATISTICS, MCMASTER UNIVERSITY, 1280 MAIN STREET WEST,
HAMILTON, ONTARIO, CANADA L8S 4K1

E-mail address: sawyer@mcmaster.ca

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING MI
E-mail address: ignacio@math.msu.edu